
WIREHOLE
***** (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Mar 8 23:35:29 1999; MasPar time 72.58 Seconds
879.119 Million cell updates/sec
Tabular output not generated.
Title: >US-09-206-040-1
Description: (1-489) from US09206040A.seq
Perfect Score: 465
N.A. Sequence: 1 ttaactgtcagcgcgcaggt.....gatgatacattgaagatnnt 469
Comp: aattgaacgtgcgcgcaggtcca.....ctactatgtaacttcanna

Scoring table: TABLE jmetric
Gap 60
Nmatch STD : Dbase 0; Query 0
Searched: 188442 seqs, 68026449 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 7.077; Variance 2.661; scale 2.659
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES							
Result	No.	Score	Match Length	DB ID	Description	Pred. No.	
1	28	6.0	570 37	T91275	cdna encoding a novel	1.36e-09	
2	28	6.0	969 38	T96721	Human synaptogyrin cd	1.36e-09	
3	28	6.0	1581 22	T32233	Cuphea C14-0-ACP thio	1.36e-09	
4	28	6.0	1698 40	V17199	Human MKK3-interactin	1.36e-09	
5	21	4.5	1257 24	T38266	54 gene differentia	1.09e-03	
6	21	4.5	2055 24	T38260	10 gene differentia	1.09e-03	
7	18	3.9	36 13	T97945	Primer SALADT contg.	2.20e-01	
8	18	3.9	787 21	T29008	Parietaria allergen p	2.20e-01	
9	18	3.9	845 21	T29009	Parietaria allergen p	2.20e-01	
10	18	3.9	1930 40	V16738	CDNA encoding human C	2.20e-01	
11	18	3.9	1930 18	T15610	CROC-1 cdna encodes c	2.20e-01	
12	18	3.9	2721 2	Q10543	BamHI J-I fragment ca	2.20e-01	
13	18	3.9	2721 2	Q10212	BamHI J-I fragment ca	2.20e-01	

14	17	3.7	26 36	T98445	Template switching ol	1.17e+00
15	17	3.7	35 13	T97346	Primer SALAD for the	1.17e+00
16	17	3.7	543 35	T39800	Human clone 34 gene.	1.17e+00
17	17	3.7	925 40	V16740	CDNA encoding human C	1.17e+00
18	17	3.7	925 18	T15611	CROC-4 cdna encodes c	1.17e+00
19	17	3.7	946 13	O80219	Human NDF-alpha3 clon	1.17e+00
20	17	3.7	1003 35	T85935	Rat epithelial membra	1.17e+00
21	17	3.7	1098 13	T80225	Rat NDF clone 4 DNA.	1.17e+00
22	17	3.7	1101 34	T74988	Interferon gamma indu	1.17e+00
23	17	3.7	1101 34	T74987	Interferon gamma indu	1.17e+00
24	17	3.7	1261 13	O80230	Rat NDF clone 40 DNA.	1.17e+00
25	17	3.7	1551 13	O80218	Human NDF-alpha2b clo	1.17e+00
26	17	3.7	1865 40	T99584	Bovine platelet activ	1.17e+00
27	17	3.7	1700 16	Q96100	Sequence encoding hum	1.17e+00
28	17	3.7	1734 35	T39792	Human SH3P9 gene.	1.17e+00
29	17	3.7	1784 38	T93211	Mouse neurotactin cdn	1.17e+00
30	17	3.7	1807 13	O80217	Human NDF-alpha2b clo	1.17e+00
31	17	3.7	1983 26	T47497	Porcine growth hormon	1.17e+00
32	17	3.7	2232 37	T95700	Novel human gene, des	1.17e+00
33	17	3.7	2288 32	T72894	Plasmodium ebl-1 gene	1.17e+00
34	17	3.7	2288 14	Q83526	P. falciparum EBL-e1	1.17e+00
35	17	3.7	2531 13	O80227	Rat NDF clone 20 DNA.	1.17e+00
36	17	3.7	2606 14	O83527	P. falciparum E31a ge	1.17e+00
37	17	3.7	2806 32	T72895	Plasmodium E31a gene.	1.17e+00
38	17	3.7	2712 24	T38265	200 gene differentia	1.17e+00
39	17	3.7	2743 13	O80231	Rat NDF clone 41 DNA.	1.17e+00
40	17	3.7	2914 13	O80226	Rat NDF clone 19 DNA.	1.17e+00
41	17	3.7	3097 31	T69590	Murine short form Ob	1.17e+00
42	17	3.7	3451 38	V02308	Cell membrane proton-	1.17e+00
43	17	3.7	3854 31	T69591	Murine long form Ob r	1.17e+00
44	17	3.7	5030 36	T75702	pGreenlntern vector.	1.17e+00
45	16	3.4	925 40	V16740	CDNA encoding human C	5.89e+00

ALIGNMENTS

RESULT 1
ID T91275 standard; cdna; 570 BP.
AC T91275;
DT 23-APR-1998 (first entry)
DE CDNA encoding a novel beta-chemokine designated PTEC.
KW Beta-chemokine; PTEC; homologue; human RANTES; C-C chemokine;
KW diagnosis; treatment; activated T cell;
KW immediate type hypersensitivity; leukocyte proliferation; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 110..391 /*tag= a
FT WO9739126-A1.
PD 23-OCT-1997.
PF 14-APR-1997; U06249.
PR 17-APR-1996; US-633682.
PA (INCY-) INCYTE PHARM INC.
PI Bandman O, Hawkins PR, Murry LE;
DR WPI; 97-526462/48.
DR P-PSDB; W27271.
PT PTEC, a novel human beta-chemokine - useful for diagnosing and
PT treating e.g. viral, bacterial, fungal infections, auto-immune
PT diseases, etc
PS Claim 2; Pages 39-40; 59pp; English.
CC The present DNA sequence encodes a novel beta-chemokine polypeptide
CC designated PTEC. The protein has Cys residues of a beta-chemokine (C32,
CC C33, C56 and C72). The present sequence was isolated from a cDNA library
CC generated from non-tumorous human prostate tissue. The PTEC protein is
CC a homologue of human RANTES, which is a C-C chemokine with structural
CC similarity to interleukin-8 and human MIP-beta. The polynucleotide
CC sequence encoding PTEC can be used to diagnose or treat a condition
CC associated with the presence of activated T cells including viral,
CC bacterial, fungal or helminthic infections, allergic or asthmatic
CC responses, mechanical injury associated with trauma, arteriosclerosis,
CC atherogenesis or collagen vascular diseases, autoimmune diseases such as
CC rheumatoid arthritis, myasthenia gravis, systemic lupus erythematosus or
CC haemolytic anaemia, leukaemia, lymphomas or carcinomas, and diseases of

CC immediate type hypersensitivity which involve activation or excessive
 CC proliferation of leukocytes, particularly monocytes, macrophages,
 CC eosinophils, basophil, mast and T cells.
 SQ Sequence 570 BP; 149 A; 155 C; 138 G; 128 T;

Query Match 6.0%; Score 28; DB 37; Length 570;
 Best Local Similarity 100.0%; Pred. No. 1.36e-09;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 59 ggaattccgggtgacccacgctccg 86
 |||||||
 Qy 29 ggaattccgggtgacccacgctccg 56

RESULT 2
 ID T96721 standard; cDNA; 969 BP.
 AC T96721; 1998 (first entry)
 DE Human synaptogyrin cDNA.
 KW Synaptogyrin; human; colon cancer; lung cancer; brain tumour;
 KW Crohn's disease; rheumatoid arthritis; AIDS; allergy; urticaria;
 KW juvenile diabetes; scleroderma; antisense; diagnosis; therapy; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 110..784
 FT /*tag= a
 PN W09741143-A1.
 PD 06-NOV-1997.
 PF 30-APR-1997; U07378.
 PR 30-APR-1996; US-700637.
 PA (INCY-) INCYTE PHARM INC.
 PI Hawkins PR, Murry LE, Stuart SG;
 DR WPI: 97-549677/50.
 DR P-PSDB; W36516.
 PT Novel human synaptogyrin homologue - useful for diagnosis and
 PT treatment of, e.g. rheumatoid arthritis, Crohn's disease, cancers,
 PT etc.
 PS Claim 2; Page 46-47; 74pp; English.
 CC This polynucleotide (snpg) identifies and encodes a novel human
 CC synaptogyrin homologue (SNPG) (see W36516). It was initially
 CC identified in Incyte Clone 775426 from a colon cDNA library
 CC (COLN005) prepared from a Crohn's disease patient. Also claimed
 CC are: (1) an expression vector containing PN; (2) a host cell
 CC transformed with the vector of (1); (3) the purified 224 amino acid
 CC SNPG protein; (4) an antisense molecule complementary to PN; (5) an
 CC antibody (Ab) specific for the polypeptide of (3); and (6) an
 CC inhibitor specific for the polypeptide of (3). Human SNPG can be
 CC used for the diagnosis and treatment of a disease associated with
 CC synaptogyrin, especially Crohn's disease, rheumatoid arthritis,
 CC asthma and cancers or tumours of the lung, colon or brain (claimed).
 CC The antisense molecule, Ab or inhibitor can be used to treat, e.g.
 CC AIDS, allergy, urticaria, juvenile diabetes, scleroderma, etc.
 SQ Sequence 969 BP; 163 A; 324 C; 280 G; 202 T;

Query Match 6.0%; Score 28; DB 38; Length 969;
 Best Local Similarity 100.0%; Pred. No. 1.36e-09;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 59 ggaattccgggtgacccacgctccg 86
 |||||||
 Qy 29 ggaattccgggtgacccacgctccg 56

RESULT 3
 ID T32233 standard; cDNA; 1581 BP.
 AC T32233;
 DT 26-OCT-1996 (first entry)
 DE Cuphea C14:0-ACP thioesterase cDNA clone MCT34 (CpFatB2).
 KW Myristate; myristic acid; fatty acid; acyl-ACP thioesterase;
 KW seed oil; oilseed; vegetable oil; Brassica; transgenic plant;
 KW surfactant; ss.
 OS Cuphea palustris.
 FH Key Location/Qualifiers

FT cds 119..1354
 FT /*tag= a
 PN W09623892-A2.
 PD 08-AUG-1996.
 PF 01-FEB-1996; U01585.
 PR 02-FEB-1995; US-383756.
 PR 05-JUN-1995; US-460898.
 PA (CALJ) CALGENE INC.
 PI Dehesh K, Hawkins D, Voelker TA;
 DR WPI: 96-371439/37.
 DR P-PSDB; W02081.
 PT Recombinant prodn. of myristate in plant cells - using DNA with
 PT preferential activity on C14 fatty acids from Cuphea palustris,
 PT nutmeg and camphor, useful in detergent and food industries
 PS Example 1; Fig 1A-1E; 77pp; English.
 CC A cDNA clone (T32233), designated MCT34 (pCpFatB2), codes for Cuphea
 CC palustris C14:0-ACP thioesterase (W02081), an enzyme that acts
 CC primarily on C14:0-ACP substrates, forming C14:0 (myristate). It
 CC was isolated from a developing seed cDNA library by screening with
 CC sequences encoding medium-chain acyl-ACP thioesterases from Cuphea
 CC hookeriana. Constructs for expression of clone MCT34 in plant seeds
 CC under the control of napin or oleosin regulatory regions were prep'd.
 CC These allow high-level prodn. of myristate (useful in surfactants
 CC and foods) in plant cells, e.g. Brassica.
 SQ Sequence 1581 BP; 418 A; 356 C; 410 G; 397 T;

Query Match 6.0%; Score 28; DB 22; Length 1581;
 Best Local Similarity 100.0%; Pred. No. 1.36e-09;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 54 ggaattccgggtgacccacgctccg 81
 |||||||
 Qy 29 ggaattccgggtgacccacgctccg 56

RESULT 4
 ID V17199 standard; cDNA; 1698 BP.
 AC V17199;
 DT 10-JUN-1998 (first entry)
 DE Human MKK3-interacting protein (MIP) encoding cDNA.
 KW MKK3-interacting protein; MIP; mitogen activated protein kinase;
 KW MAPK kinase-3; human; transduction; target; extracellular signal;
 KW interfering mutant; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 167..1213
 FT /*tag= a
 FT /product= "MKK3-interacting protein"
 PN W09800539-A2.
 PD 08-JAN-1998.
 PF 02-JUL-1997; U10866.
 PR 01-JUL-1997; US-886572.
 PR 03-JUL-1996; US-021224.
 PR 12-JUL-1996; US-021641.
 PA (CHIR) CHIRON CORP.
 PI Kavanaugh MW, Shyamala V;
 DR WPI: 98-086965/08.
 DR P-PSDB; W52169.
 PT Human MKK3-interacting protein, used to identify human tissues in
 PT malignant conditions - comprises nucleic acid, antibodies, fusion
 PT proteins and peptide fragments
 PS Disclosure: Page 14; 17pp; English.
 CC This cDNA encodes a human MKK3-interacting protein (MIP). The mitogen-
 CC activated protein kinase (MAPK) kinase-3 (MKK3) is a protein kinase that
 CC phosphorylates p38 MAP kinase specifically and is involved in transducing
 CC stress signals. The MIP binds to and interacts with a dominant
 CC interfering mutant form of MKK3 and may be involved in transducing
 CC extracellular signals to the nucleus, resulting in activation of p38
 CC kinase. MIP is expressed in human brain, kidney, liver, lung, pancreas
 CC and spleen but not in heart or striated muscle. Detection of MIP in a
 CC body sample (particularly a tumour or metastasis) is used to identify the
 CC tissue source, the tissue selected from the group consisting of brain,
 CC kidney, liver, lung, pancreas or spleen tissue. Fusion proteins comprising

CC at least 8 contiguous amino acids of MIP fused to a second protein
 CC sequence can be used to target MIP, or its peptides, to specific cells or
 CC tissues. They can be used as a source of MIP in assays and also as
 CC immunogen. The MIP encoding nucleic acid is used for expression of
 CC recombinant MIP polypeptides and as a source of probes for detecting
 CC MIP-related mRNA.

SQ Sequence 1698 BP; 455 A; 381 C; 449 G; 413 T;

Query Match 6.0%; Score 28; DB 40; Length 1698;

Best Local Similarity 100.0%; Pred. No. 1.36e-09;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 75 ggaattccgggtcgaccacgcgtccg 102

|||||

QY 29 ggaattccgggtcgaccacgcgtccg 56

RESULT

ID T38266 standard; cDNA; 1257 BP.

AC T38266;

DT 29-DEC-1996 (first entry)

DE 54 gene differentially expressed in T helper cells.

KW T helper cell; TH cell; T-cell; T-lymphocyte; 54 gene;

KW differential expression; immune disorder; multiple sclerosis;

KW asthma; lepromatous leprosy; diagnosis; therapy;

KW cysteine protease; ss.

OS Mus sp.

FS Key

FT cds Location/Qualifiers

FT 22..1137

FT /tag= a

FT mat_peptide 397..1134

FT /tag= b

FT misc_difference 1147

FT /tag= c

FT /note= "base 1147 is given as v in the

FT specification"

PN WO9627603-A1.

PD 12-SEP-1996.

PF 01-MAR-1996; U02798.

PR 03-MAR-1995; US-398633.

PR 07-JUN-1995; US-487748.

PA (MILL-) MILLENNIUM PHARM INC.

PI Levinson DA;

DR WPI: 96-433404/43.

DR P-PSDB: W01048.

PT Genes and their products differentially expressed in T helper cells

PT multiple sclerosis, asthma, lepromatous leprosy, etc.

PT Claim 1: Fig 22; 218pp; English.

PS Novel murine genes 10, 57, 105, 106, 161, 200 and 54 (T38260-66)

CC were identified in transgenic T-cell paradigm and T-cell clone

CC within and/or among T helper (TH) cell subpopulations. The 54

CC gene (T38266) is differentially expressed within the TH cell

CC subpopulation; expression levels are many-fold higher in TH cells

CC than in TH2 cells. The 54 gene codes for a novel cysteine protease

CC (W01048) that may be involved in tissue degradation. It is

CC possible that inhibition of 54 gene expression and/or 54 gene

CC product activity in immune disorders involving TH1-like cells may

CC serve to minimise tissue damage.

SQ Sequence 1257 BP; 322 A; 359 C; 322 G; 253 T;

Query Match 4.5%; Score 21; DB 24; Length 1257;

Best Local Similarity 100.0%; Pred. No. 1.09e-03;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ccgggtcgaccacgcgtccg 21

|||||

QY 36 ccgggtcgaccacgcgtccg 56

|||||

RESULT 6

ID T38260 standard; cDNA; 2055 BP.

AC T38260;

DT 29-DEC-1996 (first entry)

DE 10 gene differentially expressed in T helper cells.

KW T helper cell; TH cell; T-cell; T-lymphocyte; 10 gene;

KW differential expression; immune disorder; multiple sclerosis;

KW asthma; lepromatous leprosy; diagnosis; therapy; ss.

OS Mus sp.

FS Key

FT cds Location/Qualifiers

FT 496..1509

FT /tag= a

FT misc_feature 1106..1392

FT /tag= b

FT /note= "band 10 nucleotide sequence"

PN WO9627603-A1.

PD 12-SEP-1996.

PF 01-MAR-1996; U02798.

PR 03-MAR-1995; US-398633.

PR 07-JUN-1995; US-487748.

PA (MILL-) MILLENNIUM PHARM INC.

PI Levinson DA;

DR WPI: 96-433404/43.

DR P-PSDB: W01048.

PT Genes and their products differentially expressed in T helper cells

PT multiple sclerosis, asthma, lepromatous leprosy, etc.

PS Claim 1: Fig 9; 218pp; English.

CC Novel murine genes 10, 57, 105, 106, 161, 200 and 54 (T38260-66)

CC were identified in transgenic T-cell paradigm and T-cell clone

CC within and/or among T helper (TH) cell subpopulations. The product

CC (W01048) of the 10 gene (T38260) is a receptor protein possibly

CC involved in signal transduction. Gene 10 expression is up-regulated

CC in stimulated TH1 and TH2 cells. The gene is located on the mouse

CC chromosome 12. The gene and its product can be used to reduce the

CC level of induced T-cell activity and bring about the amelioration

CC of a wide range of T-cell related disorders.

SQ Sequence 2055 BP; 623 A; 461 C; 414 G; 557 T;

Query Match 4.5%; Score 21; DB 24; Length 2055;

Best Local Similarity 100.0%; Pred. No. 1.09e-03;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ccgggtcgaccacgcgtccg 21

|||||

QY 36 ccgggtcgaccacgcgtccg 56

|||||

RESULT 7

ID Q79345 standard; cDNA; 36 BP.

AC Q79345;

DT 05-JUN-1995 (first entry)

DE Primer SALADTT contg. a 5' SalI site, for the analysis of ARD 1.

KW ADP-ribosylation factor; ARF; ARD 1; primer; ss.

OS Synthetic.

PN WO9424283-A.

PD 27-OCT-1994.

PF 15-APR-1994; U04190.

PR 16-APR-1993; US-049252.

PR 19-APR-1993; US-049473.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Mishima K, Moss J, Nightingale M, Tsuchiya M;

DR WPI: 94-341862/42.

PT GTP-binding protein ARD1 with ADP-ribosylation factor domain -

PT useful as biochemical and diagnostic reagent

PS Example; table iii, page 16; 52pp; English.

CC A novel ARD 1 protein includes an 18 kDa region that

CC exhibits significant homology to known ADP-ribosylation factors

CC (ARFs) and is called ARD 1 for ARF domain. cDNA was isolated from a

CC human HL-60 lambda library screened with ARF 2B cDNA and a mixture of

CC oligos denoted XARFC. The insert of Clone no. 76 was sequenced. The

CC insert (bps 706-2365 Q79326) included an ORF (1207-1722) encoding

CC an ARF domain of 172 AAs. Oligo JIR was used to screen a human

CC fetal brain cDNA lambda ZAP library. Clones were found contg.

CC nucleotides 7-1826 and 726-3225 of Q79326. In this sequence, about
 CC 1200 nts preceded the ARF region without a stop codon in the same
 CC ORF. To further characterise the 5' terminus of this cDNA, 5'-RACE
 CC was carried out with the poly A+ RNA from IMR-32 human
 CC neuroblastoma cells. The poly A+ RNA was reverse transcribed with
 CC primer JK721RC. The products were tailed and subcloned to PCR with
 CC primers SALAD and SALADT and JK723RII. TE was added to the PCR
 CC products and a second PCR amplification was carried out using
 CC primers SALAD and JKNOT. Following transfection of the reverse
 CC transcribed, PCR amplified sequences into competent DH5-alpha
 CC cells, clones corresp. to ARD 1 were isolated.
 SQ Sequence 36 BP: 4 A; 12 C; 11 G; 9 T;

Query Match 3.9%; Score 18; DB 13; Length 36;
 Best Local Similarity 100.0%; Pred. No. 2.20e-01;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 ggtcgaccacgcgtccg 36
 |||||

Qy 40 gtcgaccacgcgtccg 57

RESULT 8
 ID T29008 standard; cDNA; 787 BP.

AC T29008;
 DT 16-SEP-1996 (first entry)
 DE Parietaria allergen PJ003 PAR70 cDNA (preliminary sequence).
 KW Allergen; pollen; pollinosis; desensitisation; immunotherapy;
 KW vaccine; hayfever; ds.
 OS Parietaria judaica.
 PN EP-707065-A2.
 PD 17-APR-1996.
 PF 24-JUL-1995; 401754.
 PR 22-JUL-1994; US-279113.
 PA (UYMA-) UNIV MANITOBA.
 PI Mohapatra SS;
 DR WPI; 96-201752/21.
 PT DNA encoding Parietaria sp. pollen allergen - useful in
 PT immuno-therapy and desensitisation of allergies
 PS Claim 4; Fig 3f; 34pp; English.
 CC Screening of a Parietaria pollen library with antisera to a 14 kDa
 CC pollen protein from Parietaria led to the identification of 3 groups
 CC of cDNA clones encoding allergenic proteins: group I (PJ001)
 CC comprises clones PAR64, PAR16, PAR5, PAR21, PAR48, PAR39 and
 CC PAR9 (T29000-06) encoding 7 isoallergens (R97300-06); group II (PJ009)
 CC comprises clone PAR19 (T29007) encoding 2 isoallergens (R97307-08);
 CC and group III (PJ003) comprising clones PAR70 and PAR10 (T29008-09;
 CC preliminary sequence data) with 3 isoallergenic variants. The cDNA
 CC clones can be used as probes to identify other allergen sequences, or
 CC for prodn. of recombinant allergens useful in the diagnosis and
 CC immunotherapy of allergic diseases induced by Parietaria.
 SQ Sequence 787 BP; 239 A; 189 C; 167 G; 192 T;

Query Match 3.9%; Score 18; DB 21; Length 787;
 Best Local Similarity 100.0%; Pred. No. 2.20e-01;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 ggtcgaccacgcgtccg 28
 |||||

Qy 39 ggtcgaccacgcgtccg 56

RESULT 9
 ID T29009 standard; cDNA; 845 BP.

AC T29009;
 DT 16-SEP-1996 (first entry)
 DE Parietaria allergen PJ003 PAR10 cDNA (preliminary sequence).
 KW Allergen; pollen; pollinosis; desensitisation; immunotherapy;
 KW vaccine; hayfever; ds.
 OS Parietaria judaica.
 PN EP-707065-A2.
 PD 17-APR-1996.
 PF 24-JUL-1995; 401754.

PR 22-JUL-1994; US-279113.
 PA (UYMA-) UNIV MANITOBA.
 PI Mohapatra SS;
 DR WPI; 96-201752/21.
 PT DNA encoding Parietaria sp. pollen allergen - useful in
 PT immuno-therapy and desensitisation of allergies
 PS Claim 4; Fig 3f; 34pp; English.
 CC Screening of a Parietaria pollen library with antisera to a 14 kDa
 CC pollen protein from Parietaria led to the identification of 3 groups
 CC of cDNA clones encoding allergenic proteins: group I (PJ001)
 CC comprises clones PAR64, PAR16, PAR5, PAR21, PAR48, PAR39 and
 CC PAR9 (T29000-06) encoding 7 isoallergens (R97300-06); group II (PJ009)
 CC comprises clone PAR19 (T29007) encoding 2 isoallergens (R97307-08);
 CC and group III (PJ003) comprising clones PAR70 and PAR10 (T29008-09;
 CC preliminary sequence data) with 3 isoallergenic variants. The cDNA
 CC clones can be used as probes to identify other allergen sequences, or
 CC for prodn. of recombinant allergens useful in the diagnosis and
 CC immunotherapy of allergic diseases induced by Parietaria.
 SQ Sequence 845 BP; 260 A; 194 C; 183 G; 208 T;

Query Match 3.9%; Score 18; DB 21; Length 845;
 Best Local Similarity 100.0%; Pred. No. 2.20e-01;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 ggtcgaccacgcgtccg 22
 |||||

Qy 39 ggtcgaccacgcgtccg 56

RESULT 10

ID V16738 standard; cDNA; 1930 BP.
 AC V16738;
 DT 18-JUN-1998 (first entry)
 DE cDNA encoding human CROC-1 protein.
 KW CROC-1; activating protein; promoter; proto-oncogene; c-fos;
 KW antagonist; ds.
 OS Homo sapiens.
 FH Key
 FT CDS
 FT 1..516
 FT /*tag= a
 FT /product= CROC-1
 FT US5736331-A.
 PN 07-APR-1998.
 PD 18-OCT-1995; 544900.
 PF 08-JUL-1994; US-272412.
 PR (SCHE) SCHERING CORP.
 PA Lin SL, Rothofsky ML;
 DR WPI; 98-239204/21.
 DR P-PSDB; W46892.
 PT c-fos Promoter-activating proteins - assay for nucleic acids
 PT encoding such proteins, assay for antagonists, etc.
 PS Claim 16; Columns 15-18; 12pp; English.
 CC The present sequence encodes a novel human protein, designated CROC-1.
 CC The protein is an activating protein of promoter of the proto-oncogene
 CC c-fos. The specification describes a mammalian cell line whose cells
 CC contain a recombinant expression vector comprising a reporter operatively
 CC linked to a human c-fos promoter and a second expression vector
 CC comprising a nucleic acid encoding that activates the human c-fos
 CC promoter (e.g. CROC-1). Antigenic fragments of the CROC-1 protein can be
 CC used to prepare antibodies, which are used to assay for or purify the
 CC activating protein. Antagonists of the CROC-1 protein bind to, but do not
 CC activate, the human c-fos promoter.
 SQ Sequence 1930 BP; 524 A; 487 C; 404 G; 515 T;

Query Match 3.9%; Score 18; DB 40; Length 1930;
 Best Local Similarity 100.0%; Pred. No. 2.20e-01;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 55 ggtcgaccacgcgtccg 72
 |||||

Qy 39 ggtcgaccacgcgtccg 56

```

RESULT 11
ID T15610 standard; cDNA; 1930 BP.
AC T15610;
DE 25-JUN-1996 (first entry)
DE CROC-1 cDNA encodes c-fos promoter activating protein.
KW CROC-1; CROC-4; c-fos promoter activating protein; signal transducer;
KW polyomavirus large T antigen; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT 1..516
FT /*tag= a
FT /product= CROC-1
PN W09601899-A1.
PD 25-JAN-1996.
PF 05-JUL-1995; U07874.
PR 08-JUL-1994; US-272412.
PA (SCHE ) SCHERING CORP.
PI Lin SL, Rothofsky ML;
DR WPI; 96-097629/10.
DR P-PSDB; R90769.
PT Mammalian cell lines contg. recombinant vectors encoding c-fos
PT promoter activating proteins - used to identify nucleic acids
PT encoding signal transducing molecules which activate promoters.
PS Claim 13; Page 29-31; 41pp; English.
CC CROC-1 cDNA encodes a c-fos promoter activating protein of approx. 19 kD
CC with an acidic amino terminal half and a basic carboxy terminus. The
CC protein includes a kinase target domain which contains phosphorylation
CC sites for a variety of kinases involved in signal transduction. CROC-1
CC mRNA is approx. 2.3 kb in length and present in all tissues examined.
CC The cDNA is used in vectors, operatively linked to a nucleic acid
CC encoding a polyomavirus large T antigen. Mammalian cell lines contg.
CC these vectors, and vectors comprising a polyoma origin of replication
CC and a nucleic acid suspected to encode and activating protein of the
CC promoter, can be used to identify signal transducing molecules.
SQ Sequence 1930 BP; 524 A; 487 C; 404 G; 515 T;

Query Match 3.9%; Score 18; DB 18; Length 1930;
Best Local Similarity 100.0%; Pred. No. 2.20e-01;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 55 ggtcgaccacgcgctcg 72
QY 39 ggtcgaccacgcgctcg 56

RESULT 12
ID Q10543 standard; DNA; 2721 BP.
AC Q10543;
DE 27-MAR-1991 (first entry)
DE BamHI J-I fragment carrying sequences characteristic of latent
DE pseudorabies virus.
KW PRV; SS.
OS Pseudorabies virus.
PN US7537855-A.
PD 18-DEC-1990.
PF 13-JUN-1990; 238940.
PR 13-JUN-1990; US-537855.
PA (USDA ) US AGRIC RES SERV.
PI Cheung AK;
DR WPI; 91-021957/03.
PT Pseudo-raabies nucleotide sequences - used for producing
PT nucleic acid probes, antigens and antibodies for distinguishing
PT latent from productive infection
PS Disclosure; Page 22; 27pp; English.
CC The fragment carries sequences characteristic of the latent
CC pseudorabies viral genome, and may be used as a probe in diagnosis
CC of infection.
SQ Sequence 2721 BP; 428 A; 1007 C; 1017 G; 269 T;

Query Match 3.9%; Score 18; DB 2; Length 2721;
Best Local Similarity 100.0%; Pred. No. 2.20e-01;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 55 ggtcgaccacgcgctcg 72
QY 39 ggtcgaccacgcgctcg 56

RESULT 13
ID Q10212 standard; DNA; 2721 BP.
AC Q10212;
DE 27-MAR-1991 (first entry)
DE BamHI J-I fragment carrying sequences characteristic of productive
DE pseudorabies virus.
KW PRV; SS.
OS Pseudorabies virus.
PN US7537855-A.
PD 18-DEC-1990.
PF 13-JUN-1990; 238940.
PR 13-JUN-1990; US-537855.
PA (USDA ) US AGRIC RES SERV.
PI Cheung AK;
DR WPI; 91-021957/03.
PT Pseudo-raabies virus nucleotide sequences - used for producing
PT nucleic acid probes, antigens and antibodies for distinguishing
PT latent from productive infection
PS Disclosure; Page 21; 27pp; English.
CC The fragment carries sequences characteristic of the productive
CC pseudorabies viral genome, and may be used as a probe in diagnosis
CC of infection.
SQ Sequence 2721 BP; 428 A; 1007 C; 1017 G; 269 T;

Query Match 3.9%; Score 18; DB 2; Length 2721;
Best Local Similarity 100.0%; Pred. No. 2.20e-01;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 141 ctctctttccccccccc 158
Cp 96 ctctctttccccccccc 79

RESULT 14
ID T98445 standard; DNA; 26 BP.
AC T98445;
DE 11-MAR-1998 (first entry)
DE Template switching oligonucleotide Na21-G5p.
KW Template switching oligonucleotide; RNA/DNA hybrid; DNA preparation;
KW cDNA synthesis primer; cDNA cloning; cDNA library construction; SS.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_RNA 22..26
FT /*tag= a
FT modified_base 26
FT /*tag= b
FT /*note= "phosphorylated"
PN W09724455-A2.
PD 10-JUL-1997.
PF 03-JAN-1997; U00368.
PR 03-JAN-1996; US-582562.
PA (CLON-) CLONTECH LAB INC.
PI Chenchak A, Diatchenko L, Siebert P, Zhu Y;
DR WPI; 97-363690/33.
PT Preparation of cDNA from RNA molecules - by annealing cDNA synthesis
PT primer to RNA, synthesising DNA and contacting with novel template
PT switching oligo;nucleotide
PS Claim 9; Page 29; 39pp; English.
CC T98426-T98496 represent template switching oligonucleotides (TSO) used in
CC the method of the invention. The method of the invention is for preparing
CC DNA complementary to the 5'-end of an RNA molecule. The method comprises
CC annealing a cDNA synthesis primer (such as T98497) to the RNA molecule
CC and synthesising a first DNA strand complementary to at least a portion
CC of the RNA molecule, and contacting the RNA molecule with a TSO having a
CC pre-selected nucleotide sequence at its 5'-end and at least 1 riboguanine
CC residue at its 3'-end, where the TSO binds the 5'-end of the RNA molecule
CC and serves as a template for the extension of the 3'-end of the first DNA
CC strand. The method can be used for the synthesis and cloning of full

```

CC length cDNA, or fragments, that correspond to the complete sequence of
CC the 5'-end of the mRNA molecule. It can be used to construct cDNA
CC libraries from nanogram quantities of total or poly A+ RNA. The TSO
CC allows for negative selection against cDNA that are not complementary to
CC the 5'-end of the template RNA, while allowing full length cDNA to be
CC readily selected.
SQ Sequence 26 BP; 8 A; 3 C; 12 G; 3 T;

Query Match 3.7%; Score 17; DB 36; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.17e+00;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 10 aagacgacagaaggagg 26
|||||
Qy 68 aagacgacagaaggagg 84

RESULT 15

ID Q79346 standard; cDNA; 35 BP.
AC Q79346;
DT 05-JUN-1995 (first entry)
DE Primer SALAD for the analysis of ARD 1.
KW ADP-ribosylation factor; ARF; ARD 1; Primer; ss.
OS Synthetic.
PN WO9424283-A.
PD 27-OCT-1994.
PF 15-APR-1994; U04190.
PR 16-APR-1993; US-049252.
PR 19-APR-1993; US-049473.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Mishima K, Moss J, Nightingale M, Tsuchiya M;
DR WPI; 94-341862/42.
PT GTP-binding protein ARD1 with ADP-ribosylation factor domain -
PT useful as biochemical and diagnostic reagent
PS Example; Table III, Page 16; 52pp; English.
CC A novel ARD 1 protein includes an 18 kDa region that
CC exhibits significant homology to known ADP-ribosylation factors
CC (ARFs) and is called ARD 1 for ARF domain. cDNA was isolated from a
CC human HL-60 lambda library screened with ARF 2B cDNA and a mixture of
CC oligos denoted XARFC. The insert of Clone no. 76 was sequenced. The
CC insert (bps 706-2365 Q79326) included an ORF (1207-1722) encoding
CC an ARF domain of 172 AAs. Oligo J1R was used to screen a human
CC fetal brain cDNA lambda ZAP library. Clones were found contg.
CC nucleotides 7-1826 and 726-3225 of Q79326. In this sequence, about
CC 1200 nts preceded the ARF region without a stop codon in the same
CC ORF. To further characterise the 5' terminus of this cDNA, 5'-RACE
CC was carried out with the poly A+ RNA from IMR-52 human
CC neuroblastoma cells. The poly A+ RNA was reverse transcribed with
CC primer JK721RC. The products were tailed and subcloned to PCR with
CC primers SALAD and SALADT and JK723RII. TE was added to the PCR
CC products and a second PCR amplification was carried out using
CC primers SALAD and JKNOT. Following transfection of the reverse
CC transcribed, PCR amplified sequences into competent DH5-alpha
CC cells, clones corresp. to ARD 1 were isolated.
SQ Sequence 35 BP; 4 A; 12 C; 11 G; 8 T;

Query Match 3.7%; Score 17; DB 13; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.17e+00;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 gtcgaccacgcgtccg 35
|||||
Qy 40 gtcgaccacgcgtccg 56

Search completed: Mon Mar 8 23:40:43 1999
Job time : 314 secs.

M P S R E F H

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Mar 8 23:41:01 1999; MasPar time 35.32 Seconds
Tabular output not generated. 1148.278 Million cell updates/sec

Title: >US-09-206-040-1
Description: (1-469) from US09206040A.seq
Perfect Score: 465
N.A. Sequence: 1 ttaacttcgagcgcgcaggt.....gatgatacatggaagatnnt 469
Comp: aattgaacgtcgnggtcca.....ctactatgtaactctcanna

Scoring table: TABLE jmetric
Gap 60

Nmatch STD : Dbase 0; Query 0

Searched: 165359 seqs, 43243793 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-issued
1:5A_COMB 2:5B_COMB 3:5C_COMB 4:PCT9_COMB 5:backfiles1

Statistics: Mean 6.835; Variance 2.301; scale 2.970

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.	
1	28	6.0	570	3	US-08-633- Sequence 1, Applicatio	1.25e-11	
2	28	6.0	969	3	US-08-700- Sequence 1, Applicatio	1.25e-11	
3	28	6.0	1023	3	US-08-785- Sequence 2, Applicatio	1.25e-11	
4	28	6.0	1529	3	US-08-726- Sequence 1, Applicatio	1.25e-11	
5	28	6.0	1581	2	US-08-383- Sequence 1, Applicatio	1.25e-11	
6	28	6.0	1581	3	US-08-460- Sequence 1, Applicatio	1.25e-11	
7	21	4.5	1257	2	US-08-487- Sequence 11, Applicati	5.73e-05	
8	21	4.5	2055	2	US-08-487- Sequence 3, Applicatio	5.73e-05	
9	18	3.9	36	1	US-08-049- Sequence 20, Applicati	2.33e-02	
10	18	3.9	36	4	US-08-312- Sequence 20, Applicati	2.33e-02	
11	18	3.9	36	4	PCT-US94-0 Sequence 20, Applicati	2.33e-02	
12	18	3.9	1930	4	PCT-US95-0 Sequence 1, Applicatio	2.33e-02	
13	18	3.9	1930	2	US-08-544- Sequence 1, Applicatio	2.33e-02	
14	18	3.9	2721	5	5215881-2 Patent No. 5215881	2.33e-02	
15	17	3.7	35	1	US-08-049- Sequence 21, Applicati	1.55e-01	
16	17	3.7	35	1	US-08-312- Sequence 21, Applicati	1.55e-01	
17	17	3.7	35	4	PCT-US94-0 Sequence 21, Applicati	1.55e-01	
18	17	3.7	538	3	US-08-890- Sequence 1, Applicatio	1.55e-01	
19	17	3.7	925	2	US-08-544- Sequence 3, Applicatio	1.55e-01	
20	17	3.7	925	4	PCT-US95-0 Sequence 2, Applicatio	1.55e-01	

21	17	3.7	1700	4	PCT-US95-0 Sequence 3, Applicatio	1.55e-01
22	17	3.7	1983	1	US-08-073- Sequence 9, Applicatio	1.55e-01
23	17	3.7	2288	3	US-08-568- Sequence 5, Applicatio	1.55e-01
24	17	3.7	2606	3	US-08-568- Sequence 7, Applicatio	1.55e-01
25	17	3.7	2710	2	US-08-487- Sequence 8, Applicatio	1.55e-01
26	16	3.4	16	3	US-08-407- Sequence 44, Applicati	9.74e-01
27	16	3.4	35	3	US-08-447- Sequence 77, Applicati	9.74e-01
c 28	16	3.4	39	2	US-08-846- Sequence 11, Applicati	9.74e-01
c 29	16	3.4	39	2	US-08-385- Sequence 11, Applicati	9.74e-01
c 30	16	3.4	925	4	PCT-US95-0 Sequence 2, Applicatio	9.74e-01
c 31	16	3.4	925	4	US-08-344- Sequence 3, Applicatio	9.74e-01
32	16	3.4	975	3	US-08-842- Sequence 9, Applicatio	9.74e-01
33	16	3.4	975	2	US-08-671- Sequence 9, Applicatio	9.74e-01
34	16	3.4	975	2	US-08-672- Sequence 9, Applicatio	9.74e-01
35	16	3.4	975	3	US-08-842- Sequence 9, Applicatio	9.74e-01
36	16	3.4	1264	3	US-08-758- Sequence 13, Applicatio	9.74e-01
37	16	3.4	1425	2	US-08-846- Sequence 1, Applicatio	9.74e-01
38	16	3.4	1425	2	US-08-385- Sequence 2, Applicatio	9.74e-01
39	16	3.4	1762	3	US-08-742- Sequence 1, Applicatio	9.74e-01
40	16	3.4	4010	3	US-08-785- Sequence 3, Applicatio	9.74e-01
41	16	3.4	4447	3	US-08-304- Sequence 3, Applicatio	9.74e-01
c 42	16	3.4	8540	4	PCT-US96-0 Sequence 12, Applicati	9.74e-01
c 43	16	3.4	8575	4	PCT-US92-0 Sequence 6, Applicatio	9.74e-01
c 44	16	3.4	10580	1	US-08-196- Sequence 1, Applicatio	9.74e-01
c 45	16	3.4	10596	4	PCT-US93-0 Sequence 15, Applicati	9.74e-01

ALIGNMENTS

RESULT 1
ID US-08-633-682-1 STANDARD; DNA; UNC; 570 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/08633682
CC Sequence 1, Application US/08633682
CC Patent No. 5840544
CC GENERAL INFORMATION:
CC APPLICANT: Hawkins, Phillip R.
CC APPLICANT: Bandman, Olga
CC APPLICANT: Murry, Lynn E.
CC TITLE OF INVENTION: NOVEL RANTES HOMOLOG FROM PROSTATE
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: U.S.
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq Version 1.5
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/633,682
CC FILING DATE: filed Herewith
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Luther, Barbara J
CC REGISTRATION NUMBER: 33,954
CC REFERENCE/DOCKET NUMBER: PF-0063 US
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-852-0195
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 570 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC LIBRARY: Prostate

```
CC CLONE: 836820
SQ SEQUENCE 570 BP; 149 A; 155 C; 137 G; 129 T; 0 OTHER.

Query Match 6.0%; Score 28; DB 3; Length 570;
Best Local Similarity 100.0%; Pred. No. 1.25e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 59 GGAATCCCGGTCGACCCAGCGTCG 86
| | | | | | | | | | | | | | | | | |
Qy 29 ggaattcccggtcgacccacgcgtccg 56

RESULT 2
ID US-08-700-637-1 STANDARD; DNA; UNC; 969 BP.
AC xxxxxx
DE
Dt
Sequence 1, Application US/08700637
CC Sequence 1, Application US/08700637
CC Patent No. 5854413
CC GENERAL INFORMATION:
CC APPLICANT: Hawkins, Phillip R.
CC APPLICANT: Stuart, Susan G.
CC APPLICANT: Murry, Lynn E.
CC TITLE OF INVENTION: NOVEL SYNAPTOGYRIN HOMOLOG FROM COLON
CC NUMBER OF SEQUENCES: 12
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: U.S.
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq Version 1.5
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/700.637
CC FILING DATE: Filed Herewith
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Luther, Barbara J.
CC REGISTRATION NUMBER: 33,954
CC REFERENCE/DOCKET NUMBER: PF-0065 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-852-0195
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 969 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC LIBRARY: COLNOT05
CC CLONE: 775426
CC SEQUENCE 969 BP; 163 A; 324 C; 280 G; 202 T; 0 OTHER.

Query Match 6.0%; Score 28; DB 3; Length 969;
Best Local Similarity 100.0%; Pred. No. 1.25e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 59 GGAATCCCGGTCGACCCAGCGTCG 86
| | | | | | | | | | | | | | | | | |
Qy 29 ggaattcccggtcgacccacgcgtccg 56

RESULT 3
ID US-08-785-065-2 STANDARD; DNA; UNC; 1023 BP.
AC xxxxxx
DE
Dt
Sequence 2, Application US/08785065
```

```
CC Sequence 2, Application US/08785065
CC Patent No. 5814451
CC GENERAL INFORMATION:
CC APPLICANT: Bandman, Olga
CC APPLICANT: Goli, Surya K.
CC APPLICANT: Hillman, Jennifer L.
CC TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
CC NUMBER OF SEQUENCES: 12
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/785.065
CC FILING DATE: Herewith
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Billings, Lucy J.
CC REGISTRATION NUMBER: 36,749
CC REFERENCE/DOCKET NUMBER: PF-0187 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1023 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: Consensus
CC CLONE: Consensus
CC SEQUENCE 1023 BP; 217 A; 270 C; 290 G; 246 T; 0 OTHER.

Query Match 6.0%; Score 28; DB 3; Length 1023;
Best Local Similarity 100.0%; Pred. No. 1.25e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 54 GGAATCCCGGTCGACCCAGCGTCG 81
| | | | | | | | | | | | | | | | | |
Qy 29 ggaattcccggtcgacccacgcgtccg 56

RESULT 4
ID US-08-726-575A-1 STANDARD; DNA; UNC; 1529 BP.
AC xxxxxx
DE
Dt
Sequence 1, Application US/08726575A
CC Sequence 1, Application US/08726575A
CC Patent No. 5834587
CC GENERAL INFORMATION:
CC APPLICANT: Winnie Chan, Derk J. Bergsma,
CC APPLICANT: Catherine E. Ellis
CC TITLE OF INVENTION: A NO. 5834587el G-Protein Coupled Receptor,
CC TITLE OF INVENTION: HLTEX11
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SmithKline Beecham Corporation
CC STREET: 709 Swedeland Road, P.O. Box 1539
CC CITY: King of Prussia
CC STATE: PA
```


CC COUNTRY: USA
CC ZIP: 19406-0939
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
CC COMPUTER: IBM 486
CC OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
CC SOFTWARE: WORDPERFECT 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/726,575A
CC FILING DATE: OCTOBER 8, 1996
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: William T. Han
CC REGISTRATION NUMBER: 34,344
CC REFERENCE/DOCKET NUMBER: ATG 50025
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 610 270 5219
CC TELEFAX: 610 270 4026
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1529
CC TYPE: Nucleic Acid
CC STRANDEDNESS: Single
CC TOPOLOGY: Linear
CC ANTI-SENSE: NO
CC SEQUENCE 1529 BP; 390 A; 374 C; 315 G; 450 T; 0 OTHER.
SQ
Query Match 6.0%; Score 28; DB 3; Length 1529;
Best Local Similarity 100.0%; Pred. No. 1.25e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 25 GGAATCCCGGTCGACCCACGCGTCG 52
|||||
QY 29 ggaattcccggtcgaccacggtccg 56
RESULT 5
ID US-08-383-756-1 STANDARD; DNA; UNC; 1581 BP.
AC xxxxxx
DT
Sequence 1, Application US/08383756
Sequence 1, Application US/08383756
Patent No. 5654495
CC GENERAL INFORMATION:
CC APPLICANT: Debesh, Katayoon
CC APPLICANT: Voelker, Toni Alois
CC APPLICANT: Hawkins, Deborah
CC APPLICANT: Davies, Huw Maelor
CC TITLE OF INVENTION: Production of Myristate in Plant Cells
CC NUMBER OF SEQUENCES: 17
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Calgene, Inc.
CC STREET: 1920 Fifth Street
CC CITY: Davis
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 95616
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
CC COMPUTER: Apple Macintosh
CC OPERATING SYSTEM: Macintosh 7.0
CC SOFTWARE: Microsoft Word 5.1(a)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/383,756
CC FILING DATE: 02-FEB-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/261,695
CC FILING DATE: 16-JUN-94
CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/10814
CC FILING DATE: 29-OCT-93
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: USSN 07/968,971
CC FILING DATE: 30-OCT-92
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Elizabeth Lassen
CC REGISTRATION NUMBER: 31,845
CC NAME: Donna E. Scherer
CC REGISTRATION NUMBER: 34,719
CC NAME: Carl J. Schwedler
CC REGISTRATION NUMBER: 36,924
CC REFERENCE/DOCKET NUMBER: CGNE 111
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (916) 753-6313
CC TELEFAX: (916) 753-1510
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1581 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC SEQUENCE 1581 BP; 418 A; 356 C; 410 G; 397 T; 0 OTHER.
SQ
Query Match 6.0%; Score 28; DB 2; Length 1581;
Best Local Similarity 100.0%; Pred. No. 1.25e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 54 GGAATCCCGGTCGACCCACGCGTCG 81
|||||
QY 29 ggaattcccggtcgaccacggtccg 56
RESULT 6
ID US-08-460-898-1 STANDARD; DNA; UNC; 1581 BP.
AC xxxxxx
DT
Sequence 1, Application US/08460898
Sequence 1, Application US/08460898
Patent No. 5850022
CC GENERAL INFORMATION:
CC APPLICANT: Debesh, Katayoon
CC APPLICANT: Voelker, Toni Alois
CC APPLICANT: Hawkins, Deborah
CC APPLICANT: Davies, Huw Maelor
CC TITLE OF INVENTION: Production of Myristate in Plant Cells
CC NUMBER OF SEQUENCES: 17
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Calgene, Inc.
CC STREET: 1920 Fifth Street
CC CITY: Davis
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 95616
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
CC COMPUTER: Apple Macintosh
CC OPERATING SYSTEM: Macintosh 7.0
CC SOFTWARE: Microsoft Word 5.1(a)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/460,898
CC FILING DATE: 05-JUN-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/383,756
CC FILING DATE: 02-FEB-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/261,695
CC FILING DATE: 16-JUN-94
CC CLASSIFICATION: 435

```
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/10814
CC FILING DATE: 29-OCT-93
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: USSN 07/968,971
CC FILING DATE: 30-OCT-92
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Elizabeth Lassen
CC REGISTRATION NUMBER: 31,845
CC NAME: Donna E. Scherer
CC REGISTRATION NUMBER: 34,719
CC NAME: Carl J. Schwedler
CC REGISTRATION NUMBER: 36,924
CC REFERENCE/DOCKET NUMBER: CGNE 111-1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (916) 753-6313
CC TELEFAX: (916) 753-1510
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1581 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cdna to mRNA
CC SEQUENCE 1581 BP; 418 A; 356 C; 410 G; 397 T; 0 OTHER.

Query Match 6.0%; Score 28; DB 3; Length 1581;
Best Local Similarity 100.0%; Pred.No. 1.25e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 54 GGAATCCCGGTCGACCCAGCGTCGCG 81
QY 29 ggaattcccggtcgaccacgcgctccg 56

RESULT 7
ID US-08-487-748A-11 STANDARD; DNA; UNC; 1257 BP.
AC xxxxxx
DT
DE Sequence 11, Application US/08487748A
CC Sequence 11, Application US/08487748A
CC Patent No. 5721351
CC GENERAL INFORMATION:
CC APPLICANT: Levinson, Douglas A.
CC TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
CC TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
CC NUMBER OF SEQUENCES: 23
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/487,748A
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 536
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/487,748A
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Coruzzi, Laura A.
CC REGISTRATION NUMBER: 30,742
CC REFERENCE/DOCKET NUMBER: 7853-023
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-8864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2055 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 496..1509
CC SEQUENCE 2055 BP; 623 A; 460 C; 415 G; 557 T; 0 OTHER.

Query Match 4.5%; Score 21; DB 2; Length 2055;
Best Local Similarity 100.0%; Pred.No. 5.73e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/10814
CC FILING DATE: 29-OCT-93
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: USSN 07/968,971
CC FILING DATE: 30-OCT-92
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Elizabeth Lassen
CC REGISTRATION NUMBER: 31,845
CC NAME: Donna E. Scherer
CC REGISTRATION NUMBER: 34,719
CC NAME: Carl J. Schwedler
CC REGISTRATION NUMBER: 36,924
CC REFERENCE/DOCKET NUMBER: CGNE 111-1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (916) 753-6313
CC TELEFAX: (916) 753-1510
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1581 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cdna to mRNA
CC SEQUENCE 1581 BP; 418 A; 356 C; 410 G; 397 T; 0 OTHER.

Query Match 6.0%; Score 28; DB 3; Length 1581;
Best Local Similarity 100.0%; Pred.No. 1.25e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 54 GGAATCCCGGTCGACCCAGCGTCGCG 81
QY 29 ggaattcccggtcgaccacgcgctccg 56

RESULT 7
ID US-08-487-748A-11 STANDARD; DNA; UNC; 1257 BP.
AC xxxxxx
DT
DE Sequence 11, Application US/08487748A
CC Sequence 11, Application US/08487748A
CC Patent No. 5721351
CC GENERAL INFORMATION:
CC APPLICANT: Levinson, Douglas A.
CC TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
CC TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
CC NUMBER OF SEQUENCES: 23
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/487,748A
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 536
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/487,748A
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Coruzzi, Laura A.
CC REGISTRATION NUMBER: 30,742
CC REFERENCE/DOCKET NUMBER: 7853-023
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-8864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2055 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 496..1509
CC SEQUENCE 2055 BP; 623 A; 460 C; 415 G; 557 T; 0 OTHER.

Query Match 4.5%; Score 21; DB 2; Length 2055;
Best Local Similarity 100.0%; Pred.No. 5.73e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC INFORMATION FOR SEQ ID NO: 11:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1257 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 22..1137
CC SEQUENCE 1257 BP; 323 A; 357 C; 323 G; 253 T; 1 OTHER.

Query Match 4.5%; Score 21; DB 2; Length 1257;
Best Local Similarity 100.0%; Pred.No. 5.73e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 CCGGGTCGACCCAGCGTCGCG 21
QY 36 ccgggtcgaccacgcgctccg 56

RESULT 8
ID US-08-487-748A-3 STANDARD; DNA; UNC; 2055 BP.
AC xxxxxx
DT
DE Sequence 3, Application US/08487748A
CC Sequence 3, Application US/08487748A
CC Patent No. 5721351
CC GENERAL INFORMATION:
CC APPLICANT: Levinson, Douglas A.
CC TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
CC TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
CC NUMBER OF SEQUENCES: 23
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/487,748A
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Coruzzi, Laura A.
CC REGISTRATION NUMBER: 30,742
CC REFERENCE/DOCKET NUMBER: 7853-023
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-8864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2055 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 496..1509
CC SEQUENCE 2055 BP; 623 A; 460 C; 415 G; 557 T; 0 OTHER.

Query Match 4.5%; Score 21; DB 2; Length 2055;
Best Local Similarity 100.0%; Pred.No. 5.73e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 1 CCGGTCGACCCAGCGTCGT 21
Qy 36 ccgggtcgaccacgcgtccg 56

RESULT 9
ID US-08-049-473-20 STANDARD; DNA; UNC; 36 BP.
AC xxxxxx

Sequence 20, Application US/08049473
Sequence 20, Application US/08049473
Patent No. 5385021
GENERAL INFORMATION:
CC APPLICANT: Moss, Joel
CC APPLICANT: Mishima, Koichi
CC APPLICANT: Nightingale, Maria
CC APPLICANT: Tsuchiya, Mikako
CC TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
CC TITLE OF INVENTION: PROTEIN WITH AN ADP-RIBOSYLATION FACTOR DOMAIN
CC NUMBER OF SEQUENCES: 34
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
CC STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CC CITY: NEWPORT BEACH
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 92660

COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/049,473
CC FILING DATE: 19930419
CC CLASSIFICATION: 436
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Fuller, Michael L.
CC REGISTRATION NUMBER: 36,516
CC REFERENCE/DOCKET NUMBER: NIH050.001CP1
CC TELEPHONE: 619-235-8550
CC TELEFAX: 619-235-0176
CC INFORMATION FOR SEQ ID NO: 20:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 36 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC SEQUENCE 36 BP; 4 A; 12 C; 11 G; 9 T; 0 OTHER.

Query Match 3.9%; Score 18; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 2.33e-02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 GTCGACCCAGCGTCGT 36
Qy 40 gtcgacccacgcgtccgt 57

RESULT 10
ID US-08-312-648-20 STANDARD; DNA; UNC; 36 BP.
AC xxxxxx

Sequence 20, Application US/08312648
Sequence 20, Application US/08312648
Patent No. 5514600
GENERAL INFORMATION:
CC APPLICANT: Moss, Joel
CC APPLICANT: Mishima, Koichi

CC APPLICANT: Nightingale, Maria
CC APPLICANT: Tsuchiya, Mikako
CC TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
CC TITLE OF INVENTION: PROTEIN WITH AN ADP-RIBOSYLATION FACTOR DOMAIN
CC NUMBER OF SEQUENCES: 34
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
CC STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CC CITY: NEWPORT BEACH
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 92660

COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/312,648
CC FILING DATE:
CC CLASSIFICATION: 436
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/049,473
CC FILING DATE: 19-APR-1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Fuller, Michael L.
CC REGISTRATION NUMBER: 36,516
CC REFERENCE/DOCKET NUMBER: NIH050.001DV1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619-235-8550
CC TELEFAX: 619-235-0176
CC INFORMATION FOR SEQ ID NO: 20:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 36 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC SEQUENCE 36 BP; 4 A; 12 C; 11 G; 9 T; 0 OTHER.

Query Match 3.9%; Score 18; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 2.33e-02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 GTCGACCCAGCGTCGT 36
Qy 40 gtcgacccacgcgtccgt 57

RESULT 11
ID PCT-US94-04190-20 STANDARD; DNA; UNC; 36 BP.
AC xxxxxx

Sequence 20, Application PC/TUS9404190
Sequence 20, Application PC/TUS9404190
GENERAL INFORMATION:
CC APPLICANT: The Government of the United States of America
CC APPLICANT: as represented by the Secretary, Department
CC APPLICANT: of Health and Human Services
CC TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
CC TITLE OF INVENTION: PROTEIN WITH AN ADP-RIBOSYLATION FACTOR DOMAIN
CC NUMBER OF SEQUENCES: 34
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
CC STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CC CITY: NEWPORT BEACH
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 92660

COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/04190
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Fuller, Michael L.
CC REGISTRATION NUMBER: 36,516
CC REFERENCE/DOCKET NUMBER: NIH050.001QPC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619-235-8550
CC TELEFAX: 619-235-0176
CC INFORMATION FOR SEQ ID NO: 20:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 36 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC SEQUENCE 36 BP; 4 A; 12 C; 11 G; 9 T; 0 OTHER.

Query Match 3.9%; Score 18; DB 4; Length 36;
Best Local Similarity 100.0%; Pred. No. 2.33e-02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 GTCGACCCACGCGTCGT 36
Qy 40 gtcgacccacgcgtcgt 57

RESULT 12
ID PCT-US95-07874-1 STANDARD; DNA; UNC; 1930 BP.
AC xxxxxx

DE Sequence 1, Application PC/TUS9507874
CC Sequence 1, Application PC/TUS9507874
CC GENERAL INFORMATION:

CC APPLICANT:
CC TITLE OF INVENTION: Method for Identifying Nucleic Acids Encoding c-fos Promoter
CC NUMBER OF SEQUENCES: 2
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: Apple Macintosh
CC OPERATING SYSTEM: Macintosh 7.1
CC SOFTWARE: Microsoft Word 5.1a
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/07874
CC FILING DATE: June-1995

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/272,412
CC FILING DATE: 8-JUL-1994

CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1930 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear

CC SEQUENCE 1930 BP; 524 A; 487 C; 404 G; 515 T; 0 OTHER.
Query Match 3.9%; Score 18; DB 4; Length 1930;
Best Local Similarity 100.0%; Pred. No. 2.33e-02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 55 GTCGACCCACGCGTCGT 72
Qy 39 gtcgacccacgcgtcgcg 56

RESULT 13

ID US-08-544-900-1 STANDARD; DNA; UNC; 1930 BP.
AC xxxxxx
DE Sequence 1, Application US/08544900
CC Sequence 1, Application US/08544900
CC Patent No. 5736331
CC GENERAL INFORMATION:
CC APPLICANT: Lin, Stanley Li
CC APPLICANT: Rothofsky, Marnie Lynn
CC TITLE OF INVENTION: Method for Identifying Nucleic
CC TITLE OF INVENTION: Acids Encoding c-fos Promoter
CC NUMBER OF SEQUENCES: 3
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Schering-Plough Corporation
CC ADDRESSEE: Patent Department K-6-1 (1990)
CC STREET: 2000 Galloping Hill Road
CC CITY: Kenilworth
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 07033-0530
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: Apple Macintosh
CC OPERATING SYSTEM: Macintosh 7.1
CC SOFTWARE: Microsoft Word 5.1a
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/544,900
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/272,412
CC FILING DATE: 8-JUL-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Thompson, Paul A.
CC REGISTRATION NUMBER: 35,385
CC REFERENCE/DOCKET NUMBER: OC0439K
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 908 298 5150
CC TELEFAX: 908 298 5388
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1930 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC SEQUENCE 1930 BP; 524 A; 487 C; 404 G; 515 T; 0 OTHER.

Query Match 3.9%; Score 18; DB 2; Length 1930;
Best Local Similarity 100.0%; Pred. No. 2.33e-02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 55 GTCGACCCACGCGTCGT 72
Qy 39 gtcgacccacgcgtcgcg 56

RESULT 14
ID 5215881-2 STANDARD; DNA; UNC; 2948 BP.

AC xxxxxx
DE 01-JAN-1900
DE Patent No. 5215881.

CC Patent No. 5215881
CC APPLICANT: CHEUNG, ANDREW K.
CC TITLE OF INVENTION: PSEUDORABIES DIAGNOSIS PROBES
CC NUMBER OF SEQUENCES: 3
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/537,855
CC FILING DATE: 13-JUN-1990
CC SEQ ID NO:2:

```
CC LENGTH: 2721
SQ Sequence 2948 BP; 428 A; 1007 C; 1017 G; 269 T; 227 other;

Query Match          3.9%; Score 18; DB 5; Length 2721;
Best Local Similarity 100.0%; Pred. No. 2.33e-02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 141 CTCCTTTTCCCCCCT 158
Cp 96 ctctcttccccccct 79

RESULT 15
ID US-08-049-473-21 STANDARD; DNA; UNC; 35 BP.
AC xxxxxx
DT
DE Sequence 21, Application US/08049473
CC Sequence 21, Application US/08049473
CC Patent No. 5386021
CC GENERAL INFORMATION:
CC APPLICANT: Moss, Joel
CC APPLICANT: Mishima, Koichi
CC APPLICANT: Nightingale, Maria
CC APPLICANT: Tsuchiya, Mikako
CC TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
CC TITLE OF INVENTION: PROTEIN WITH AN ADP-RYBOSYLATION FACTOR DOMAIN
CC NUMBER OF SEQUENCES: 34
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
CC STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CC CITY: NEWPORT BEACH
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 92660
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/049,473
CC FILING DATE: 19930419
CC CLASSIFICATION: 436
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Fuller, Michael L.
CC REGISTRATION NUMBER: 36,516
CC REFERENCE/DOCKET NUMBER: NIH050.001CPI
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619-235-8550
CC TELEFAX: 619-235-0176
CC INFORMATION FOR SEQ ID NO: 21:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 35 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
SQ SEQUENCE 35 BP; 4 A; 12 C; 11 G; 8 T; 0 OTHER.

Query Match          3.7%; Score 17; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.55e-01;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 GTCGACCCACGCGTCG 35
Qy 40 gtcgaccacgcgctccg 56

Search completed: Mon Mar 8 23:42:51 1999
Job time : 110 secs.
```

THIS PAGE BLANK (USPTO)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Mon Mar 8 23:10:19 1999; MasPar time 622.20 seconds
 Tabular output not generated. 1349.844 Million cell updates/sec

Title: >US-09-206-040-1
 Description: (1-469) from US09206040A.seq
 Perfect Score: 465
 N.A. Sequence: 1 ttaacttgacgcnccaggt.....gatgatacattgaagatntt 469
 Comp: attggaactgcgnggtccca.....ctactatgtaacttctanna

Scoring table: TABLE jmetric
 Gap 60

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: emb1-est56

Database: genbank-est109
 5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
 10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
 14:gb_est18 15:gb_est19 16:gb_est2 17:gb_est20
 18:gb_est21 19:gb_est3 20:gb_est4 21:gb_est5 22:gb_est6
 23:gb_est7 24:gb_est8 25:gb_est9 26:gb_gss1 27:gb_gss2
 28:gb_gss3 29:gb_gss4

Statistics: Mean 8.493; Variance 1.124; scale 7.553

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	30	6.5	568	21	5C04F08-T7 membrane-fr	3.59e-31
2	29	6.2	214	21	5C06D02-T7 membrane-fr	8.16e-29
3	29	6.2	331	21	5C06H03-T7 membrane-fr	8.16e-29
4	29	6.2	407	21	5C06A03-T7 membrane-fr	8.16e-29
5	29	6.2	439	21	5C06G03-T7 membrane-fr	8.16e-29
6	29	6.2	589	9	Homo sapiens mRNA; exp	8.16e-29
7	28	6.0	113	21	5C01E04-T7 membrane-fr	1.72e-26
8	28	6.0	114	21	5C06H05-T7 membrane-fr	1.72e-26
9	28	6.0	163	21	5C04E02-T7 membrane-fr	1.72e-26
10	28	6.0	184	21	5C06C10-T7 membrane-fr	1.72e-26
11	28	6.0	221	21	5C04F02-T7 membrane-fr	1.72e-26
12	28	6.0	340	23	Human clone xs359 mRNA	1.72e-26
13	28	6.0	350	23	Human clone xs360 mRNA	1.72e-26

14	28	6.0	360	23	HSU46350	1.72e-26
15	28	6.0	386	21	5C04G09-T7 membrane-fr	1.72e-26
16	28	6.0	397	11	2742 Lambda-PRL2 Arabi	1.72e-26
17	28	6.0	400	23	Human Clone xs363 mRNA	1.72e-26
18	28	6.0	416	21	5C04G04-T7 membrane-fr	1.72e-26
19	28	6.0	419	21	5C06G08-T7 membrane-fr	1.72e-26
20	28	6.0	430	23	Human clone xs353 mRNA	1.72e-26
21	28	6.0	431	21	5C06D04-T7 membrane-fr	1.72e-26
22	28	6.0	432	21	5C01F06-T7 membrane-fr	1.72e-26
23	28	6.0	432	21	5C04A03-T7 membrane-fr	1.72e-26
24	28	6.0	436	21	5C04F09-T7 membrane-fr	1.72e-26
25	28	6.0	450	21	5C06G10-T7 membrane-fr	1.72e-26
26	28	6.0	450	21	5C06C02-T7 membrane-fr	1.72e-26
27	28	6.0	450	21	5C01H05-T7 membrane-fr	1.72e-26
28	28	6.0	450	21	5C06A01-T7 membrane-fr	1.72e-26
29	28	6.0	450	21	5C06C08-T7 membrane-fr	1.72e-26
30	28	6.0	450	21	5C06B09-T7 membrane-fr	1.72e-26
31	28	6.0	450	21	5C06G10-T7 membrane-fr	1.72e-26
32	28	6.0	450	21	5C06A09-T7 membrane-fr	1.72e-26
33	28	6.0	450	21	5C01F07-T7 membrane-fr	1.72e-26
34	28	6.0	450	21	5C06B08-T7 membrane-fr	1.72e-26
35	28	6.0	450	21	5C06A11-T7 membrane-fr	1.72e-26
36	28	6.0	450	21	5C06B02-T7 membrane-fr	1.72e-26
37	28	6.0	450	21	5C01G12-T7 membrane-fr	1.72e-26
38	28	6.0	450	21	5C01H06-T7 membrane-fr	1.72e-26
39	28	6.0	450	21	5C01H10-T7 membrane-fr	1.72e-26
40	28	6.0	450	21	5C06A08-T7 membrane-fr	1.72e-26
41	28	6.0	450	21	5C06F11-T7 membrane-fr	1.72e-26
42	28	6.0	551	23	24 Pinus radiata somat	1.72e-26
43	28	6.0	881	18	GDM00L03 human embryo	1.72e-26
44	28	6.0			GDM00L05 human embryo	1.72e-26
45	28	6.0	895	18	AI142134	1.72e-26

ALIGNMENTS

1 RESULT 1 T18698 568 bp mRNA EST 17-OCT-1996
 LOCUS 5C04F08-T7 membrane-free polysomes from endosperm Zea mays cDNA
 DEFINITION clone 5C04F08 5' end similar to 22kd alpha zein alpha precursor,
 mRNA sequence.
 ACCESSION T18698
 NID 9485628
 KEYWORDS EST.
 SOURCE maize.
 ORGANISM Zea mays.
 Eukaryotes; Mitochondrial eukaryotes; Viridiplantae;
 Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
 Liliopsida; Poales; Poaceae; Zea.
 REFERENCE 1 (bases 1 to 568)
 AUTHORS Shen,B., Carneiro,N., Torres-Jerez,I., Stevenson,R., McCreery,T.,
 Helentjaris,T., Baysdorfer,C., Almira,E., Perl,R., Habben,J. and
 Larkins,B.
 TITLE Partial sequencing and mapping of clones from two maize cDNA
 libraries
 JOURNAL Plant Mol. Biol. 26, 1085-1101 (1994)
 MEDLINE 95111093
 COMMENT Contact: The Maize cDNA Project

Helentjaris TG (primary contact)

Dept. of Plant Sciences

University of Arizona

Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721

ph: 602-6218-746

fax: 602-621-7186

E-mail: helnjars@ccit.arizona.edu

Chris Baysdorfer

Department of Biological Sciences, School of Science

California State University, Hayward

Hayward, CA 94542

ph: 510-881-3459
 fax: 510-727-2035
 E-mail: cbaysdor@sl.csuhayward.edu

Rob Ferl
 Interdisciplinary Center for Biotechnology Research
 DNA Sequencing Core
 University of Florida
 P.O. Box 100695
 Gainesville, FL 32611-0695
 ph: 904-392-1928, ext. 301
 fax: 904-392-4072
 E-mail: robferl@nervm.nerdc.ufl.edu

```

FEATURES             Location/Qualifiers
  source              1..568
                     /organism="Zea mays"
                     /strain="W64A2"
                     /note="Vector: ZipLox; Site_1: SalI; Site_2: NotI; ds-cDNA
was prepared from oligo-dr selected mRNA by priming with a
NotI oligo- dr oligomer and then adding the second strand
to RNase-nicked DNA:RNA hybrid with DNA Pol. SalI
adaptors were added to the ends, the ds-cDNAs were then
digested with NotI and size-selected. These were
directionally-cloned into the ZipLox phage vector, excised
as plasmids, and then individually analyzed."
                     /db_xref="taxon:4577"
                     /clone="5C04F08"
                     /clone_lib="membrane-free polysomes from endospERM"
                     /lab_host="DH10B"
                     />368
BASE COUNT           153 a 183 c 104 g 123 t 5 others
ORIGIN
Query Match          6.5%; Score 30; DB 21; Length 568;
Best Local Similarity 100.0%; Pred. No. 3,598-31;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 25 GGAATTCGGGTGCGACCCACGCGTCGCTA 54
|||||
Oy 29 ggaattccgggtcgaccacgcggtccgta 58

```

```

RESULT 2
LOCUS   T18296 214 bp mRNA EST 17-OCT-1996
DEFINITION 5c06d02-t7 membrane-free polysomes from endospERM Zea mays cDNA
clone 5c06d02 5' end, mRNA sequence.
ACCESSION T18296
NID 9483329
KEYWORDS EST.
SOURCE zeamays.
ORGANISM Zea mays
          Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
          Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
          Liliopsida; Poales; Poaceae; Zea.
REFERENCE 1 (bases 1 to 214)
AUTHORS Shen,B., Carneiro,N., Torres-Jerez,I., Stevenson,R., McCreery,T.,
Helentjaris,T., Baysdorfer,C., Almira,E., Ferl,R., Habben,J. and
Larkins,B.
TITLE Partial sequencing and mapping of clones from two maize cDNA
libraries
JOURNAL Plant Mol. Biol. 26, 1085-1101 (1994)
MEDLINE 95111093
COMMENT
Contact: The Maize cDNA Project

```

Helentjaris TG (primary contact)
 Dept. of Plant Sciences
 University of Arizona
 Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721

ph: 602-6218-746
 fax: 602-621-1786
 E-mail: helntjars@ccit.arizona.edu

Chris Baysdorfer
 Department of Biological Sciences, School of Science
 California State University, Hayward
 Hayward, CA 94542
 ph: 510-881-3459
 fax: 510-727-2035
 E-mail: cbaysdor@sl.csuhayward.edu

Rob Ferl
 Interdisciplinary Center for Biotechnology Research
 DNA Sequencing Core
 University of Florida
 P.O. Box 100695
 Gainesville, FL 32611-0695
 ph: 904-392-1928, ext. 301
 fax: 904-392-4072
 E-mail: robferl@nervm.nerdc.ufl.edu

```

Seq primer: T7.
Location/Qualifiers
  1..214
  /organism="Zea mays"
  /strain="W64A2"
  /note="Vector: ZipLox; Site_1: SalI; Site_2: NotI; ds-cDNA
was prepared from oligo-dr selected mRNA by priming with a
NotI oligo- dr oligomer and then adding the second strand
to RNase-nicked DNA:RNA hybrid with DNA Pol. SalI
adaptors were added to the ends, the ds-cDNAs were then
digested with NotI and size-selected. These were
directionally-cloned into the ZipLox phage vector, excised
as plasmids, and then individually analyzed."
  /db_xref="taxon:4577"
  /clone="5C06d02"
  /clone_lib="membrane-free polysomes from endospERM"
  /lab_host="DH10B"

```

```

mRNA
BASE COUNT          65 a 49 c 35 g 57 t 8 others
ORIGIN
<1..>214

```

```

Query Match          6.2%; Score 29; DB 21; Length 214;
Best Local Similarity 100.0%; Pred. No. 8,168-29;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 GGAATTCGGGTGCGACCCACGCGTCGCT 35
|||||
Oy 29 ggaattccgggtcgaccacgcggtccgt 57

```

```

RESULT 3
LOCUS   T18319 331 bp mRNA EST 17-OCT-1996
DEFINITION 5c06h03-t7 membrane-free polysomes from endospERM Zea mays cDNA
clone 5c06h03 5' end similar to ribosomal protein S8, mRNA
sequence.
ACCESSION T18319
NID 9463352
KEYWORDS EST.
SOURCE zeamays.
ORGANISM Zea mays
          Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
          Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
          Liliopsida; Poales; Poaceae; Zea.
REFERENCE 1 (bases 1 to 331)
AUTHORS Shen,B., Carneiro,N., Torres-Jerez,I., Stevenson,R., McCreery,T.,
Helentjaris,T., Baysdorfer,C., Almira,E., Ferl,R., Habben,J. and
Larkins,B.
TITLE Partial sequencing and mapping of clones from two maize cDNA
libraries
JOURNAL Plant Mol. Biol. 26, 1085-1101 (1994)

```


MEDLINE 95111093
COMMENT

Contact: The Maize CDNA Project

Helentjaris TG (Primary contact)
Dept. of Plant Sciences
University of Arizona
Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721
ph: 602-6218-746
fax: 602-621-7186
E-mail: helentjaris@ci.azizona.edu

Chris Baysdorfer
Department of Biological Sciences, School of Science
California State University, Hayward
Hayward, CA 94542
ph: 510-881-3459
fax: 510-727-2035
E-mail: cbaysdorfer@sl.csu Hayward.edu

Rob Ferl
Interdisciplinary Center for Biotechnology Research
DNA Sequencing Core
University of Florida
P.O. Box 100695
Gainesville, FL 32611-0695
ph: 904-392-1928, ext. 301
fax: 904-392-4072
E-mail: robferl@nervm.nerdc.ufl.edu

FEATURES
source

Seq primer: T7.
Location/Qualifiers
1..331
/organism="Zea mays"
/strain="W64A2"
/note="Vector: Ziplox; Site.1: SalI; Site.2: NotI; ds-cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo- dT oligomer and then adding the second strand to RNase-nicked DNA:RNA hybrid with DNA PolI. SalI adaptors were added to the ends, the ds-cDNAs were then digested with NotI and size-selected. These were directionally-cloned into the Ziplox phage vector, excised as plasmids, and then individually analyzed."
/db_xref="taxon:4577"
/clone="5c06h03"
/clone_lib="membrane-free polysomes from endosperm"
/lab_host="DH10B"
<1..>331

BASE COUNT 81 a 78 c 103 g 57 t 12 others
ORIGIN

Query Match 6.2%; Score 29; DB 21; Length 331;
Best Local Similarity 100.0%; Pred. No. 8.16e-29;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 GGAATCCCGGGTCGACCCAGCGTCGCT 37
QY 29 ggaattcccggtgcgaccagcgctcgt 57

RESULT 4
LOCUS T18280 407 bp mRNA EST 17-OCT-1996
DEFINITION 5c06a03-t7 membrane-free polysomes from endosperm Zea mays cDNA
ACCESSION T18280
NID 9463313
KEYWORDS EST.
SOURCE maize.
ORGANISM Zea mays
Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;

REFERENCE
AUTHORS

Lilipalsida; Poales; Poaceae; Zea.
1 (bases 1 to 407)

TITLE

JOURNAL
MEDLINE
95111093

COMMENT
Contact: The Maize CDNA Project

Helentjaris TG (Primary contact)
Dept. of Plant Sciences
University of Arizona
Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721
ph: 602-6218-746
fax: 602-621-7186
E-mail: helentjaris@ci.azizona.edu

Chris Baysdorfer
Department of Biological Sciences, School of Science
California State University, Hayward
Hayward, CA 94542
ph: 510-881-3459
fax: 510-727-2035
E-mail: cbaysdorfer@sl.csu Hayward.edu

Rob Ferl
Interdisciplinary Center for Biotechnology Research
DNA Sequencing Core
University of Florida
P.O. Box 100695
Gainesville, FL 32611-0695
ph: 904-392-1928, ext. 301
fax: 904-392-4072
E-mail: robferl@nervm.nerdc.ufl.edu

FEATURES
source

Seq primer: T7.
Location/Qualifiers
1..407
/organism="Zea mays"
/strain="W64A2"
/note="Vector: Ziplox; Site.1: SalI; Site.2: NotI; ds-cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo- dT oligomer and then adding the second strand to RNase-nicked DNA:RNA hybrid with DNA PolI. SalI adaptors were added to the ends, the ds-cDNAs were then digested with NotI and size-selected. These were directionally-cloned into the Ziplox phage vector, excised as plasmids, and then individually analyzed."
/db_xref="taxon:4577"
/clone="5c06a03"
/clone_lib="membrane-free polysomes from endosperm"
/lab_host="DH10B"
<1..>407

BASE COUNT 113 a 80 c 95 g 119 t
ORIGIN

Query Match 6.2%; Score 29; DB 21; Length 407;
Best Local Similarity 100.0%; Pred. No. 8.16e-29;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 GGAATCCCGGGTCGACCCAGCGTCGCT 35
QY 29 ggaattcccggtgcgaccagcgctcgt 57

RESULT 5
LOCUS T18310 439 bp mRNA EST 17-OCT-1996
DEFINITION 5c06g03-t7 membrane-free polysomes from endosperm Zea mays cDNA

accession clone 5c06g03 5' end, mRNA sequence.
 T18310
 NID 9463343
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays
 Eukaryotes; mitochondrial eukaryotes; Viridiplantae;
 Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
 Liliopsida; Poales; Poaceae; Zea.
 1 (bases 1 to 439)
 REFERENCE Shen,B., Carneiro,N., Torres-Jerez,I., Stevenson,R., McCreery,T.,
 Helentjaris,T., Baysdorfer,C., Almira,E., Ferl,R., Habbem,J. and
 Larkins,B.
 Partial sequencing and mapping of clones from two maize CDNA
 libraries
 JOURNAL Plant Mol. Biol. 26, 1085-1101 (1994)
 MEDLINE 95111093
 COMMENT Contact: The Maize cDNA Project

Helentjaris TG (primary contact)
 Dept. of Plant Sciences
 University of Arizona
 Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721
 ph: 602-6218-746
 fax: 602-621-7186
 E-mail: helentjaris@cit.arizona.edu
 Chris Baysdorfer
 Department of Biological Sciences, School of Science
 California State University, Hayward
 Hayward, CA 94542
 ph: 510-881-3459
 fax: 510-727-2035
 E-mail: cbaysdorfer1.csuhayward.edu
 Rob Ferl
 Interdisciplinary Center for Biotechnology Research
 DNA Sequencing Core
 University of Florida
 P.O. Box 100695
 Gainesville, FL 32611-0695
 ph: 904-392-1928, ext. 301
 fax: 904-392-4072
 E-mail: roberferl@nerdc.ufl.edu

FEATURES
 source

Seq primer: T7.
 Location/Qualifiers
 1..439
 /organism="Zea mays"
 /strain="M64A2"
 /note="Vector: Ziplox; Site_1: SalI; Site_2: NotI; ds-cDNA
 was prepared from oligo-dT selected mRNA by priming with a
 NotI oligo- dT oligomer and then adding the second strand
 to RNase-nicked DNA:RNA hybrid with DNA Pol. I. SalI
 adaptors were added to the ends, the ds-cDNAs were then
 digested with NotI and size-selected. These were
 directionally-cloned into the Ziplox phage vector, excised
 as plasmids, and then individually analyzed."
 /db_xref="taxon:4577"
 /clone="5c06g03"
 /clone.lib="membrane-free polysomes from endosperm"
 /lab.host="DH10B"
 <1..>439
 108 a 123 c 101 g 96 t 11 others
 MRNA
 BASE COUNT
 ORIGIN

Query Match 6.2% Score 29; DB 21; Length 439;
 Best Local Similarity 100.0%; Pred. No. 8.16e-29;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 13 GGAATTCGGGTCGACCCGCGTCGT 41
 ||||||||||||||||||||||||||||
 Oy 29 ggaatcccggtcgaccacgcgctcgt 57
 RESULT 6
 LOCUS HS299396 589 bp DNA EST 19-SEP-1997
 DEFINITION Homo sapiens mRNA; expressed sequence tag; clone DKFZphamY1_1c11,
 3' read, mRNA sequence.
 ACCESSION 299396
 NID g2415636
 KEYWORDS EST; expressed sequence tag.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Homnidae; Homo.
 REFERENCE Korn,B., Wleemann,S., Ebert,L. and Poustka,A.
 1 (bases 1 to 589)
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 589)
 AUTHORS Korn,B.
 TITLE Direct Submission
 JOURNAL Submitted (11-SEP-1997) Korn B., Institution Molekulare
 Genomanalyse, Deutsches Krebsforschungszentrum Heidelberg, Im
 Neuenheimer Feld 506, D-69120 Heidelberg, FRG
 FEATURES
 source
 1..589
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone.lib="DKFZphamY1_1c11"
 /clone="DKFZphamY1_1c11"
 /tissue_type="amygdala"
 /dev_stage="adult"
 BASE COUNT 161 a 156 c 151 g 120 t 1 others
 ORIGIN
 Query Match 6.2% Score 29; DB 9; Length 589;
 Best Local Similarity 100.0%; Pred. No. 8.16e-29;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 398 ACGGACGCGTGGTCGACCCGCGAATTC 426
 ||||||||||||||||||||||||||||
 Cp 57 agcagcgctgtgtcgaccgcggaattcc 29
 RESULT 7
 LOCUS T18257 113 bp mRNA EST 17-OCT-1996
 DEFINITION 5c01e04-t7 membrane-free polysomes from endosperm Zea mays CDNA
 clone 5c01e04 5' end, mRNA sequence.
 ACCESSION T18257
 NID 9463280
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays
 Eukaryotes; mitochondrial eukaryotes; Viridiplantae;
 Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
 Liliopsida; Poales; Poaceae; Zea.
 1 (bases 1 to 113)
 REFERENCE Shen,B., Carneiro,N., Torres-Jerez,I., Stevenson,R., McCreery,T.,
 Helentjaris,T., Baysdorfer,C., Almira,E., Ferl,R., Habbem,J. and
 Larkins,B.
 Partial sequencing and mapping of clones from two maize CDNA
 libraries
 JOURNAL Plant Mol. Biol. 26, 1085-1101 (1994)
 MEDLINE 95111093
 COMMENT Contact: The Maize cDNA Project

Helentjaris TG (primary contact)
 Dept. of Plant Sciences
 University of Arizona
 Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721

Ph: 602-6218-746
 Fax: 602-621-7186
 E-mail: helnjars@ccit.arizona.edu

Chris Baysdorfer
 Department of Biological Sciences, School of Science
 California State University, Hayward
 Hayward, CA 94542
 Ph: 510-881-3459
 Fax: 510-727-2035
 E-mail: cbaysdor@sl.csuhayward.edu

Rob Ferl
 Interdisciplinary Center for Biotechnology Research
 DNA Sequencing Core
 University of Florida
 P.O. Box 100695
 Gainesville, FL 32611-0695
 Ph: 904-392-1928, ext. 301
 Fax: 904-392-4072
 E-mail: robferl@nervm.nerdc.ufl.edu

FEATURES
 SOURCE

Seq primer: T7.
 Location/Qualifiers

1. .113
 /organism="Zea mays"
 /strain="W64A2"
 /note="Vector: Ziplox; Site_1: SalI; Site_2: NotI; ds-cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT oligomer and then adding the second strand to RNase-nicked DNA:RNA hybrid with DNA Pol. SalI adaptors were added to the ends, the ds-cDNAs were then digested with NotI and size-selected. These were directionally-cloned into the Ziplox phage vector, excised as plasmids, and then individually analyzed."
 /db_xref="taxon:4577"
 /clone="5c01e04"
 /clone_1lb="membrane-free polysomes from endosperm"
 /lab_host="DH10B"
 <1. .>113

BASE COUNT
 ORIGIN

mRNA
 29 a 29 c 33 g 22 t

Query Match 6.0%; Score 28; DB 21; Length 113;
 Best Local Similarity 100.0%; Pred. No. 1,72e-26;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 12 GGAATTCGGGTCGACCGCGTCG 39
 |||
 29 ggaattccgggtcgaccacgcgtccg 56

RESULT 8
 LOCUS T18320 114 bp mRNA EST 17-OCT-1996
 DEFINITION 5c06h05-t7 membrane-free polysomes from endosperm Zea mays cDNA
 clone 5c06h05 5' end, mRNA sequence.
 ACCESSION T18320
 NID 9463353
 KEYWORDS EST.
 SOURCE maize.
 ORGANISM Zea mays

REFERENCE
 AUTHORS Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
 Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
 Liliopsida; Poales; Poaceae; Zea.
 1 (bases 1 to 114)
 Shen, B., Carneiro, N., Torres-Jerez, I., Stevenson, R., McCreary, T.,
 Helentjaris, T., Baysdorfer, C., Almirante, E., Ferl, R., Habben, J., and
 Larkins, B.
 TITLE Partial sequencing and mapping of clones from two maize cDNA
 libraries
 JOURNAL Plant Mol. Biol. 26, 1085-1101 (1994)
 MEDLINE 95111093

COMMENT

Contact: The Maize cDNA Project

Helentjaris TG (primary contact)
 Dept. of Plant Sciences
 University of Arizona
 Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721
 Ph: 602-6218-746
 Fax: 602-621-7186
 E-mail: helnjars@ccit.arizona.edu

Chris Baysdorfer
 Department of Biological Sciences, School of Science
 California State University, Hayward
 Hayward, CA 94542
 Ph: 510-881-3459
 Fax: 510-727-2035
 E-mail: cbaysdor@sl.csuhayward.edu

Rob Ferl
 Interdisciplinary Center for Biotechnology Research
 DNA Sequencing Core
 University of Florida
 P.O. Box 100695
 Gainesville, FL 32611-0695
 Ph: 904-392-1928, ext. 301
 Fax: 904-392-4072
 E-mail: robferl@nervm.nerdc.ufl.edu

FEATURES
 SOURCE

Seq primer: T7.
 Location/Qualifiers

1. .114
 /organism="Zea mays"
 /strain="W64A2"
 /note="Vector: Ziplox; Site_1: SalI; Site_2: NotI; ds-cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT oligomer and then adding the second strand to RNase-nicked DNA:RNA hybrid with DNA Pol. SalI adaptors were added to the ends, the ds-cDNAs were then digested with NotI and size-selected. These were directionally-cloned into the Ziplox phage vector, excised as plasmids, and then individually analyzed."
 /db_xref="taxon:4577"
 /clone="5c06h05"
 /clone_1lb="membrane-free polysomes from endosperm"
 /lab_host="DH10B"
 <1. .>114

BASE COUNT
 ORIGIN

mRNA
 24 a 33 c 27 g 28 t 2 others

Query Match 6.0%; Score 28; DB 21; Length 114;
 Best Local Similarity 100.0%; Pred. No. 1,72e-26;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 10 GGAATTCGGGTCGACCGCGTCG 37
 |||
 29 ggaattccgggtcgaccacgcgtccg 56

RESULT 9
 LOCUS T18683 163 bp mRNA EST 17-OCT-1996
 DEFINITION 5C04E02-T7 membrane-free polysomes from endosperm Zea mays cDNA
 clone 5C04E02 5' end, mRNA sequence.
 ACCESSION T18683
 NID 9485613
 KEYWORDS EST.
 SOURCE maize.
 ORGANISM Zea mays

REFERENCE
 AUTHORS Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
 Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
 Liliopsida; Poales; Poaceae; Zea.

REFERENCE 1 (bases 1 to 163)
 AUTHORS Shen,B., Carneiro,N., Torres-Jerez,I., Stevenson,R., McCreery,T., Helentjaris,T., Baysdorfer,C., Almira,E., Ferl,R., Habbem,J. and Larkins,B.
 TITLE Partial sequencing and mapping of clones from two maize CDNA libraries
 JOURNAL Plant Mol. Biol. 26, 1085-1101 (1994)
 MEDLINE 95111093
 COMMENT Contact: The Maize CDNA Project

Helentjaris TG (primary contact)
 Dept. of Plant Sciences
 University of Arizona
 Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721
 ph: 602-6218-746
 fax: 602-621-7186
 E-mail: helnjars@cclt.arizona.edu

Chris Baysdorfer
 Department of Biological Sciences, School of Science
 California State University, Hayward
 Hayward, CA 94542
 ph: 510-881-3459
 fax: 510-727-2035
 E-mail: cbaysdor@sl.csu Hayward.edu

Rob Ferl
 Interdisciplinary Center for Biotechnology Research
 DNA Sequencing Core
 University of Florida
 P.O. Box 100695
 Gainesville, FL 32611-0695
 ph: 904-392-1928, ext. 301
 fax: 904-392-4072
 E-mail: robferl@nervm.nerdc.ufl.edu

FEATURES
 source
 Seq primer: T7,
 Location/Qualifiers
 1..163
 /organism="Zea mays"
 /strain="W64A2"
 /note="Vector: ziplox; Site.1: SalI; Site.2: NotI; ds-cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT oligomer and then adding the second strand to RNase-nicked DNA:RNA hybrid with DNA PolI. SalI adaptors were added to the ends, the ds-cDNAs were then digested with NotI and size-selected. These were directionally-cloned into the ziplox phage vector, excised as plasmids, and then individually analyzed."
 /db_xref="taxon:4577"
 /db_xref="taxon:4577"
 /clone_lib="membrane-free polysomes from endosperm"
 /lab_host="DH10B"
 <1..>163
 92 a 23 c 27 g 16 t 5 others

BASE COUNT
 ORIGIN
 mRNA
 BASE COUNT
 ORIGIN
 Query Match 6.0%; Score 28; DB 21; Length 163;
 Best Local Similarity 100.0%; Pred. No. 1.72e-26;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 24 GGAATTCGCGGGTCGACCGCGCTCCG 51
 ||||||||||||||||||||||||||||
 QY 29 ggaattccgggtcgacccacgcgtccg 56

RESULT 10
 LOCUS T18294 184 bp mRNA EST 17-OCT-1996
 DEFINITION Sc06c10-t7 membrane-free polysomes from endosperm Zea mays cDNA
 clone Sc06c10 5' end, mRNA sequence.

ACCESSION T18294
 NID 9463327
 KEYWORDS EST.
 SOURCE maize.
 ORGANISM Zea mays
 Eukaryote; mitochondrial eukaryotes; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
 1 (bases 1 to 184)
 AUTHORS Shen,B., Carneiro,N., Torres-Jerez,I., Stevenson,R., McCreery,T., Helentjaris,T., Baysdorfer,C., Almira,E., Ferl,R., Habbem,J. and Larkins,B.
 TITLE Partial sequencing and mapping of clones from two maize CDNA libraries
 JOURNAL Plant Mol. Biol. 26, 1085-1101 (1994)
 MEDLINE 95111093
 COMMENT Contact: The Maize CDNA Project

Helentjaris TG (primary contact)
 Dept. of Plant Sciences
 University of Arizona
 Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721
 ph: 602-6218-746
 fax: 602-621-7186
 E-mail: helnjars@cclt.arizona.edu

Chris Baysdorfer
 Department of Biological Sciences, School of Science
 California State University, Hayward
 Hayward, CA 94542
 ph: 510-881-3459
 fax: 510-727-2035
 E-mail: cbaysdor@sl.csu Hayward.edu

Rob Ferl
 Interdisciplinary Center for Biotechnology Research
 DNA Sequencing Core
 University of Florida
 P.O. Box 100695
 Gainesville, FL 32611-0695
 ph: 904-392-1928, ext. 301
 fax: 904-392-4072
 E-mail: robferl@nervm.nerdc.ufl.edu

FEATURES
 source
 Seq primer: T7,
 Location/Qualifiers
 1..184
 /organism="Zea mays"
 /strain="W64A2"
 /note="Vector: ziplox; Site.1: SalI; Site.2: NotI; ds-cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT oligomer and then adding the second strand to RNase-nicked DNA:RNA hybrid with DNA PolI. SalI adaptors were added to the ends, the ds-cDNAs were then digested with NotI and size-selected. These were directionally-cloned into the ziplox phage vector, excised as plasmids, and then individually analyzed."
 /db_xref="taxon:4577"
 /db_xref="taxon:4577"
 /clone_lib="membrane-free polysomes from endosperm"
 /clone="Sc06c10"
 /lab_host="DH10B"
 <1..>184
 42 a 38 c 42 g 61 t 1 others

BASE COUNT
 ORIGIN
 mRNA
 BASE COUNT
 ORIGIN
 Query Match 6.0%; Score 28; DB 21; Length 184;
 Best Local Similarity 100.0%; Pred. No. 1.72e-26;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 10 GGAATTCGCGGGTCGACCGCGCTCCG 37

BASE COUNT	47 a	56 c	66 g	25 t	17 others
ORIGIN					
Query Match	6.0%: Score 28; DB 21: Length 221;				
Best Local Similarity	100.0%; Pred. No. 1.72e-26;				
Matches	28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Db	25 GGAATTCGGGTCGACCCACGGTCG 52				
Qy	29 ggaattccgggtcgacccacggctcg 56				
RESULT 12					
LOCUS	HSU46345 340 bp mRNA				
DEFINITION	Human clone xs359 mRNA sequence.				
ACCESSION	U46345				
NID	91236487				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 340)				
AUTHORS	Gress, T.M., Muller-Pillasch, F., Geng, M., Zimmerhackl, F., Zehetner, G., Friess, H., Buchler, M., Adler, G. and Lehrach, H.				
TITLE	A pancreatic cancer-specific expression profile				
JOURNAL	Oncogene 13 (8), 1819-1830 (1996)				
MEDLINE	97050791				
REFERENCE	2 (bases 1 to 340)				
AUTHORS	Mueller-Pillasch, F., Zimmerhackl, F., Geng, M. and Gress, T.M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-JAN-1996) Internal Medicine I, University of Ulm, Robert-Koch-Strasse 8, Ulm, Baden-Wuerttemberg 89081, Germany				
FEATURES	Location/Qualifiers				
source	1..340				
	/organism="Homo sapiens"				
	/note="overexpressed in pancreatic cancer"				
	/db_xref="taxon:9606"				
	/cell_line="Patu 8988t"				
	/tissue_type="pancreatic cancer"				
	/clone="xs359"				
BASE COUNT	78 a	79 c	64 g	95 t	24 others
ORIGIN					
Query Match	6.0%: Score 28; DB 23: Length 340;				
Best Local Similarity	100.0%; Pred. No. 1.72e-26;				
Matches	28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Db	31 GGAATTCGGGTCGACCCACGGTCG 58				
Qy	29 ggaattccgggtcgacccacggctcg 56				
RESULT 13					
LOCUS	HSU46346 350 bp mRNA				
DEFINITION	Human clone xs360 mRNA sequence.				
ACCESSION	U46346				
NID	91236488				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 350)				
AUTHORS	Gress, T.M., Muller-Pillasch, F., Geng, M., Zimmerhackl, F., Zehetner, G., Friess, H., Buchler, M., Adler, G. and Lehrach, H.				
TITLE	A pancreatic cancer-specific expression profile				
JOURNAL	Oncogene 13 (8), 1819-1830 (1996)				
MEDLINE	97050791				
REFERENCE	2 (bases 1 to 350)				
AUTHORS	Mueller-Pillasch, F., Zimmerhackl, F., Geng, M. and Gress, T.M.				
TITLE	Direct Submission				

JOURNAL Submitted (16-JAN-1996) Internal Medicine I, University of Ulm,
Robert-Koch-Strasse 8, Ulm, Baden-Wuerttemberg 89081, Germany

FEATURES
Source
Location/Qualifiers
1..350
/organism="Homo sapiens"
/note="overexpressed in pancreatic cancer"
/db_xref="taxon:9606"
/cell_line="Patu 8988t"
/tissue_type="pancreatic cancer"
/clone="xs360"

BASE COUNT 104 a 59 c 77 g 99 t 11 others
ORIGIN

Query Match 6.0%; Score 28; DB 23; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.72e-26;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 34 GGAATTCCTCCGGTGCACCCACGCGTCCG 61
|||||
Qy 29 ggaattcccggtgcacccacgcgctccg 56

RESULT 14
LOCUS HSU46350 360 bp mRNA EST 13-MAR-1997
DEFINITION Human clone xs366 mRNA sequence.
ACCESSION U46350
NID g1236492
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 360)
Gress,T.M., Muller-Pillasch,F., Geng,M., Zimmerhackl,F.,
Zenetner,G., Friess,H., Buchler,M., Adler,G. and Lehrach,H.
A pancreatic cancer-specific expression profile
Oncogene 13 (8), 1819-1830 (1996)
97050791
2 (bases 1 to 360)
Muller-Pillasch,F., Zimmerhackl,F., Geng,M. and Gress,T.M.
Direct Submission
Submitted (16-JAN-1996) Internal Medicine I, University of Ulm,
Robert-Koch-Strasse 8, Ulm, Baden-Wuerttemberg 89081, Germany

FEATURES
Source
Location/Qualifiers
1..360
/organism="Homo sapiens"
/note="overexpressed in pancreatic cancer"
/db_xref="taxon:9606"
/cell_line="Patu 8988t"
/tissue_type="pancreatic cancer"
/clone="xs366"

BASE COUNT 53 a 95 c 129 g 67 t 16 others
ORIGIN

Query Match 6.0%; Score 28; DB 23; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.72e-26;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 40 GGAATTCCTCCGGTGCACCCACGCGTCCG 67
|||||
Qy 29 ggaattcccggtgcacccacgcgctccg 56

RESULT 15
LOCUS T18713 386 bp mRNA EST 17-OCT-1996
DEFINITION 5C04G09-T7 membrane-free polysomes from endosperm Zea mays cDNA
clone 5C04G09 5' end similar to 40s ribosomal protein S8, mRNA
sequence.
ACCESSION T18713
NID g485643
KEYWORDS EST.
SOURCE maize.
ORGANISM Zea mays

Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
Charophyta;Embryophyta group; Embryophyta; Magnoliophyta;
Lilopsida; Poales; Poaceae; Zea.
1 (bases 1 to 386)
Shen,B., Carneiro,N., Torres-Jerez,I., Stevenson,R., McCreery,T.,
Helentjaris,T., Baysdorfer,C., Almira,E., Ferl,R., Habben,J. and
Larkins,B.
Partial sequencing and mapping of clones from two maize cDNA
libraries
Plant Mol. Biol. 26, 1085-1101 (1994)
95111093

Contact: The Maize cDNA Project

Helentjaris TG (primary contact)
Dept. of Plant Sciences
University of Arizona
Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721
ph: 602-6218-746
fax: 602-621-7186
E-mail: helnjars@ccit.arizona.edu

Chris Baysdorfer
Department of Biological Sciences, School of Science
California State University, Hayward
Hayward, CA 94542
ph: 510-881-3459
fax: 510-727-2035
E-mail: cbaysdor@sl.csu Hayward.edu

Rob Ferl
Interdisciplinary Center for Biotechnology Research
DNA Sequencing Core
University of Florida
P.O. Box 100695
Gainesville, FL 32611-0695
ph: 904-392-1928, ext. 301
fax: 904-392-4072
E-mail: robferl@nervm.nerdc.ufl.edu

Seq primer: T7.
Location/Qualifiers
1..386
/organism="Zea mays"
/strain="W64A2"
/note="Vector: zipLox; Site_1: SalI; Site_2: NotI; ds-cDNA
was prepared from oligo-dt selected mRNA by priming with a
NotI oligo- dt oligomer and then adding the second strand
to RNase-nicked DNA:RNA hybrid with DNA POLI. SalI
adaptors were added to the ends, the ds-cDNAs were then
digested with NotI and size-selected. These were
directionally-cloned into the zipLox phage vector, excised
as plasmids, and then individually analyzed."
/db_xref="taxon:4577"
/clone="5C04G09"
/clone_lib="membrane-free polysomes from endosperm"
/lab_host="DH10B"
<1..>386

FEATURES
Source
Location/Qualifiers
1..386
/organism="Zea mays"
/strain="W64A2"
/note="Vector: zipLox; Site_1: SalI; Site_2: NotI; ds-cDNA
was prepared from oligo-dt selected mRNA by priming with a
NotI oligo- dt oligomer and then adding the second strand
to RNase-nicked DNA:RNA hybrid with DNA POLI. SalI
adaptors were added to the ends, the ds-cDNAs were then
digested with NotI and size-selected. These were
directionally-cloned into the zipLox phage vector, excised
as plasmids, and then individually analyzed."
/db_xref="taxon:4577"
/clone="5C04G09"
/clone_lib="membrane-free polysomes from endosperm"
/lab_host="DH10B"
<1..>386

BASE COUNT 96 a 92 c 120 g 74 t 4 others
ORIGIN

Query Match 6.0%; Score 28; DB 21; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.72e-26;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 25 GGAATTCCTCCGGTGCACCCACGCGTCCG 52
|||||
Qy 29 ggaattcccggtgcacccacgcgctccg 56

Search completed: Mon Mar 8 23:35:10 1999

Job time : 1491 secs.

PAGE BLANK (USPTO)

M P S R C H
***** (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Mar 8 22:14:44 1999; MasPar time 35.64 Seconds
Tabular output not generated.
1138.185 Million cell updates/sec

Title: >US-09-206-040-1
Perfect Score: 465
N.A. Sequence: 1 ttaactgcagcgcgcaggt.....gatgatacatgtgaagatnnt 469
Comp: aattgaagcgcgcgcaggtcca.....ctactatgaactctctanna

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 165359 seqs, 43243793 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-issued
1:5A_COMB 2:5B_COMB 3:5C_COMB 4:PCT9_COMB 5:backfiles1

Statistics: Mean 7.884; Variance 3.978; scale 1.982

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match	Length	ID	Description	Pred. No.	
1	45	9.7	7218	2	US-08-232-Sequence 14, Applicati	5.20e-17	
2	40	8.6	1023	3	US-08-785-Sequence 2, Applicatio	1.91e-13	
3	38	8.2	570	3	US-08-633-Sequence 1, Applicatio	4.81e-12	
4	38	8.2	969	3	US-08-700-Sequence 1, Applicatio	4.81e-12	
5	38	8.2	1529	3	US-08-726-Sequence 1, Applicatio	4.81e-12	
6	38	8.2	1581	3	US-08-460-Sequence 1, Applicatio	4.81e-12	
7	38	8.2	1581	2	US-08-383-Sequence 1, Applicatio	4.81e-12	
8	37	8.0	965	3	US-08-388-Sequence 22, Applicati	2.38e-11	
9	36	7.7	7218	2	US-08-232-Sequence 14, Applicati	1.16e-10	
10	29	6.2	215	1	US-08-238-Sequence 5, Applicatio	5.54e-06	
11	27	5.8	215	1	US-08-238-Sequence 5, Applicatio	1.05e-04	
12	27	5.8	538	3	US-08-890-Sequence 1, Applicatio	1.05e-04	
13	27	5.8	2606	3	US-08-568-Sequence 22, Applicatio	1.05e-04	
14	26	5.6	965	3	US-08-388-Sequence 7, Applicatio	4.45e-04	
15	25	5.4	74	4	PCT-US95-1 Sequence 92, Applicati	1.85e-03	
16	25	5.4	81	4	PCT-US95-1 Sequence 94, Applicati	1.85e-03	
17	24	5.2	74	4	PCT-US95-1 Sequence 100, Applicat	7.50e-03	
18	24	5.2	81	4	PCT-US95-1 Sequence 98, Applicati	7.50e-03	
19	23	4.9	75	4	PCT-US95-1 Sequence 99, Applicati	2.97e-02	
20	23	4.9	82	4	PCT-US95-1 Sequence 97, Applicati	2.97e-02	

21	22	4.7	74	4	PCT-US95-1	Sequence 100, Applicat	1.15e-01
c 22	22	4.7	75	4	PCT-US95-1	Sequence 99, Applicati	1.15e-01
23	22	4.7	81	4	PCT-US95-1	Sequence 98, Applicati	1.15e-01
c 24	22	4.7	81	4	PCT-US95-1	Sequence 92, Applicati	1.15e-01
c 25	22	4.7	82	4	PCT-US95-1	Sequence 97, Applicati	1.15e-01
26	22	4.7	92	3	US-08-353-Sequence 16, Applicati	1.15e-01	
27	21	4.5	42	3	US-08-452-Sequence 43, Applicati	4.31e-01	
c 28	21	4.5	54	3	US-08-452-Sequence 14, Applicati	4.31e-01	
c 29	21	4.5	54	3	US-08-452-Sequence 15, Applicati	4.31e-01	
c 30	21	4.5	65	1	US-08-471-Sequence 145, Applicat	4.31e-01	
c 31	21	4.5	66	1	US-08-471-Sequence 144, Applicat	4.31e-01	
c 32	21	4.5	68	1	US-08-471-Sequence 143, Applicat	4.31e-01	
c 33	21	4.5	69	1	US-08-471-Sequence 142, Applicat	4.31e-01	
c 34	21	4.5	74	4	PCT-US95-1	Sequence 94, Applicati	4.31e-01
35	21	4.5	163	4	PCT-US95-0	Sequence 8, Applicatio	4.31e-01
36	21	4.5	1257	2	US-08-487-Sequence 11, Applicati	4.31e-01	
37	21	4.5	1930	4	PCT-US95-0	Sequence 1, Applicatio	4.31e-01
38	21	4.5	1930	2	US-08-344-Sequence 1, Applicatio	4.31e-01	
39	21	4.5	2055	2	US-08-487-Sequence 3, Applicatio	4.31e-01	
40	20	4.3	66	1	US-08-471-Sequence 144, Applicat	1.57e+00	
41	20	4.3	69	1	US-08-471-Sequence 142, Applicat	1.57e+00	
c 42	20	4.3	105	1	US-07-865-Sequence 13, Applicati	1.57e+00	
c 43	20	4.3	242	2	US-08-273-Sequence 1, Applicatio	1.57e+00	
44	20	4.3	900	4	PCT-US95-0	Sequence 1, Applicatio	1.57e+00
45	20	4.3	24979	4	PCT-US93-0	Sequence 3, Applicatio	1.57e+00

ALIGNMENTS

RESULT 1
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC xxxxxx
DT
DE Sequence 14, Application US/08232463
CC Sequence 14, Application US/08232463
CC Patent No. 5670367
CC GENERAL INFORMATION:
CC APPLICANT: DORNER, F.
CC APPLICANT: SCHEIFLINGER, F.
CC APPLICANT: FALKNER, F. G.
CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22313-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/232,463
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/07/935,313
CC FILING DATE:
CC APPLICATION NUMBER: EP 91 114 300.6
CC FILING DATE: 26-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703)836-9300
CC TELEFAX: (703)683-4109
CC TELEX: 899149
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:

```
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: pTgpt-F1s
SQ SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

Query Match 9.7%; Score 45; DB 2; Length 7218;
Best Local Similarity 0.5%; Pred. No. 5.20e-17;
Matches 1; Conservative 121; Mismatches 83; Indels 0; Gaps 0;

Db 1225 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1284
QY 150 cttgcttgcattcgaattcattcccaattctctgcattcacaatac 209

Db 1285 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1344
QY 210 ataactgtcctaagattgcattgatttcaaaatgcagatcagggatcgagt 269

Db 1345 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1404
QY 270 cacagactcctcagatgggaataatgcacccgattcccgcgctctctcctccatg 329

Db 1405 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1435
QY 330 ttgcgcacttgcctctctatctacgtcgctg 360

RESULT 2
ID US-08-785-065-2 STANDARD; DNA; UNC; 1023 BP.
AC xxxxxx
DE
DT
Sequence 2, Application US/08785065
Sequence 2, Application US/08785065
Patent No. 5814451
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/785.065
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0187 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1023 base pairs
TYPE: nucleic acid

CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: pTgpt-F1s
SQ SEQUENCE 1023 BP; 217 A; 270 C; 290 G; 246 T; 0 OTHER.

Query Match 8.6%; Score 40; DB 3; Length 1023;
Best Local Similarity 86.5%; Pred. No. 1.91e-13;
Matches 45; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 30 CTGTAGCTGCTAGTACCGGTCCGGAATCCCGGGTCGACCCACGCGTCCG 81
QY 5 cttgcagcgnccaggtancggtcaggatcccggtcgacccacgctccg 56

RESULT 3
ID US-08-633-682-1 STANDARD; DNA; UNC; 570 BP.
AC xxxxxx
DE
DT
Sequence 1, Application US/08633682
Sequence 1, Application US/08633682
Patent No. 5840544
GENERAL INFORMATION:
APPLICANT: Hawkins, Phillip R.
APPLICANT: Bandman, Olga
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: NOVEL RANTES HOMOLOG FROM PROSTATE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633,682
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Lather, Barbara J.
REGISTRATION NUMBER: 33,954
REFERENCE/DOCKET NUMBER: PF-0063 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 570 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY: Prostrate
CLONE: 836820
SQ SEQUENCE 570 BP; 149 A; 155 C; 137 G; 129 T; 0 OTHER.

Query Match 8.2%; Score 38; DB 3; Length 570;
Best Local Similarity 95.1%; Pred. No. 4.81e-12;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 46 CAGGTACCGTCCGGAATCCCGGGTCGACCCACGCGTCCG 86
QY 16 caggtancggtcaggatcccggtcgacccacgctccg 56

RESULT 4
```

```
ID US-08-700-637-1 STANDARD; DNA; UNC; 969 BP.
AC xxxxxx
DE Sequence 1, Application US/08700637
DE Sequence 1, Application US/08700637
CC Patent No. 5854413
CC GENERAL INFORMATION:
CC APPLICANT: Hawkins, Phillip R.
CC APPLICANT: Stuart, Susan G.
CC APPLICANT: Murry, Lynn E.
CC TITLE OF INVENTION: NOVEL SYNAPTOGYRIN HOMOLOG FROM COLON
CC NUMBER OF SEQUENCES: 12
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: U.S.
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq Version 1.5
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/700,637
CC FILING DATE: Filed Herewith
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Luther, Barbara J.
CC REGISTRATION NUMBER: 33,954
CC REFERENCE/DOCKET NUMBER: PF-0065 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-852-0195
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 969 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC LIBRARY: COLNOT05
CC CLONE: 775426
CC SEQUENCE 969 BP: 163 A: 324 C: 280 G: 202 T: 0 OTHER.

Query Match 8.2%; Score 38; DB 3; Length 969;
Best Local Similarity 95.1%; Pred. No. 4.81e-12;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 46 CAGGTACCGGTCCGGAATTCCTCCGGTCCGACCCACGCGTCGCG 86
QY 16 cagggtancggtcagggaattcccggtccgacccacgcgctccg 56

RESULT 5
ID US-08-726-575A-1 STANDARD; DNA; UNC; 1529 BP.
AC xxxxxx
DE Sequence 1, Application US/08726575A
DE Sequence 1, Application US/08726575A
CC Patent No. 5834587
CC GENERAL INFORMATION:
CC APPLICANT: Winnie Chan, Derk J. Bergsma,
CC APPLICANT: Catherine E. Ellis
CC TITLE OF INVENTION: A No. 5834587el G-Protein Coupled Receptor,
CC TITLE OF INVENTION: HLTEX11
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SmithKline Beecham Corporation
CC STREET: 709 Swedeland Road, P.O. Box 1539
CC CITY: King of Prussia
CC STATE: PA

Query Match 8.2%; Score 38; DB 3; Length 969;
Best Local Similarity 95.1%; Pred. No. 4.81e-12;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 46 CAGGTACCGGTCCGGAATTCCTCCGGTCCGACCCACGCGTCGCG 86
QY 16 cagggtancggtcagggaattcccggtccgacccacgcgctccg 56

RESULT 5
ID US-08-726-575A-1 STANDARD; DNA; UNC; 1529 BP.
AC xxxxxx
DE Sequence 1, Application US/08726575A
DE Sequence 1, Application US/08726575A
CC Patent No. 5834587
CC GENERAL INFORMATION:
CC APPLICANT: Dehesh, Katayoon
CC APPLICANT: Voelker, Toni Alois
CC APPLICANT: Hawkins, Deborah
CC APPLICANT: Davies, Huw Maelor
CC TITLE OF INVENTION: Production of Myristate in Plant Cells
CC NUMBER OF SEQUENCES: 17
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Calgene, Inc.
CC STREET: 1920 Fifth Street
CC CITY: Davis
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 95616
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
CC COMPUTER: Apple Macintosh
CC OPERATING SYSTEM: Macintosh 7.0
CC SOFTWARE: Microsoft Word 5.1(a)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/460,898
CC FILING DATE: 05-JUN-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/383,756
CC FILING DATE: 02-FEB-1995
CC CLASSIFICATION: 435

Country: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,575A
FILING DATE: OCTOBER 8, 1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: William T. Han
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: ATG 50025
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610 270 5219
TELEFAX: 610 270 4026
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1529
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: NO
SEQUENCE 1529 BP: 390 A: 374 C: 315 G: 450 T: 0 OTHER.

Query Match 8.2%; Score 38; DB 3; Length 1529;
Best Local Similarity 95.1%; Pred. No. 4.81e-12;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 12 CAGGTACCGGTCCGGAATTCCTCCGGTCCGACCCACGCGTCGCG 52
QY 16 cagggtancggtcagggaattcccggtccgacccacgcgctccg 56

RESULT 6
ID US-08-460-898-1 STANDARD; DNA; UNC; 1581 BP.
AC xxxxxx
DE Sequence 1, Application US/08460898
DE Sequence 1, Application US/08460898
CC Patent No. 5850022
CC GENERAL INFORMATION:
CC APPLICANT: Dehesh, Katayoon
CC APPLICANT: Voelker, Toni Alois
CC APPLICANT: Hawkins, Deborah
CC APPLICANT: Davies, Huw Maelor
CC TITLE OF INVENTION: Production of Myristate in Plant Cells
CC NUMBER OF SEQUENCES: 17
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Calgene, Inc.
CC STREET: 1920 Fifth Street
CC CITY: Davis
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 95616
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
CC COMPUTER: Apple Macintosh
CC OPERATING SYSTEM: Macintosh 7.0
CC SOFTWARE: Microsoft Word 5.1(a)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/460,898
CC FILING DATE: 05-JUN-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/383,756
CC FILING DATE: 02-FEB-1995
CC CLASSIFICATION: 435
```


CC	REFERENCE/DOCKET NUMBER:	LUD 54.09
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE:	212-688-9200
CC	TELEFAX:	212-838-3884
CC	INFORMATION FOR SEQ ID NO:	22:
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH:	965 base pairs
CC	TYPE:	nucleic acid
CC	STRANDEDNESS:	unknown
CC	TOPOLOGY:	unknown
CC	MOLECULE TYPE:	DNA (genomic)
CC	SEQUENCE 965 BP:	192 A: 170 C: 226 G: 20

RESULT	9
ID	US-08-232-463-14 STANDARD; DNA: UNC: 7218 BP.

```

CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: PTZgpt-Fls
CC SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.
SQ
Query Match 7.78; Score 36; DB 2; Length 7218;
Best Local Similarity 4.38; Pred. No. 1.10e-10;
Matches 4; Conservative 61; Mismatches 29; Indels 0
Db 1345 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cc 149 tccttggtgcgtgcgtgcactttctgtcgaagaagtcaccaagaatccactc
Db 1405 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTAC 1438
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cc 89 tccccccctctgtctgtcttcgcagcgctac 56
RESULT 10
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DT
DE Sequence 5, Application US/08238163
CC Sequence 5, Application US/08238163
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: STOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Khourie and Crew
CC STREET: Steuart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc_feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
CC SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.
SQ

```

Query Match 6.2%; Score 29; DB 1; Length 215;
Best Local Similarity 12.5%; pred. No. 5.54e-06;
Matches 19; Conservative 65; Mismatches 67; Indels


```

RESULT 13
ID US-08-568-459A-7 STANDARD; DNA; UNC; 2606 BP.
AC xxxxxx
DE
DT
Sequence 7, Application US/08568459A
CC Sequence 7, Application US/08568459A
CC Patent No. 5849306
CC GENERAL INFORMATION:
CC APPLICANT: SIm, Kim L.
CC APPLICANT: Chitnis, Chetan
CC APPLICANT: Miller, Louis H.
CC APPLICANT: Peterson, David S.
CC APPLICANT: Su, Xin-zhaun
CC APPLICANT: Wellens, Thomas E.
CC TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
CC TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
CC NUMBER OF SEQUENCES: 37
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Knobb Martens Olson & Bear
CC STREET: 620 Newport Center Drive 16th Floor
CC CITY: Newport Beach
CC STATE: California
CC COUNTRY: US
CC ZIP: 92660
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/568,459A
CC FILING DATE: 07-DEC-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Israelsen, Ned
CC REGISTRATION NUMBER: 29,655
CC REFERENCE/DOCKET NUMBER: NIH121.001CP1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 235-8550
CC TELEFAX: (619) 235-0176
CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2606 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC HYPOTHETICAL: NO
CC ORIGINAL SOURCE:
CC ORGANISM: Plasmodium falciparum
CC SEQUENCE 2606 BP; 992 A; 351 C; 569 G; 694 T; 0 OTHER.

Query Match 5.8%; Score 27; DB 3; Length 2606;
Best Local Similarity 93.3%; Pred. No. 1.05e-04;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 42 CAGGTACCGTCCGGAAATCCCGGGTCGAC 71
QY 16 cagggtancggtcagggaattcccggtcgac 45
||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
ID US-08-388-672A-22 STANDARD; DNA; UNC; 965 BP.
AC xxxxxx
DE
DT
Sequence 22, Application US/08388672A
CC Sequence 22, Application US/08388672A
CC Patent No. 5795961
CC GENERAL INFORMATION:
CC APPLICANT: Wallace, T. Paul
CC APPLICANT: Harris, William J.
CC APPLICANT: Carr, Frank J.

```

```

CC APPLICANT: Old, Lloyd J.
CC APPLICANT: Welt, Sydney
CC APPLICANT: Kitamura, Kunio
CC TITLE OF INVENTION: Recombinant Human Anti-Lewis B
CC TITLE OF INVENTION: Antibodies
CC NUMBER OF SEQUENCES: 25
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Felfe and Lynch
CC STREET: 805 Third Avenue
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10022
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/388,672A
CC FILING DATE: 14-FEB-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hanson, No. 5795961man D.
CC REGISTRATION NUMBER: 30,946
CC REFERENCE/DOCKET NUMBER: LUD 5409
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212-688-9200
CC TELEFAX: 212-838-3884
CC INFORMATION FOR SEQ ID NO: 22:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 965 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: unknown
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 965 BP; 192 A; 170 C; 226 G; 200 T; 177 OTHER.

Query Match 5.6%; Score 26; DB 3; Length 965;
Best Local Similarity 15.8%; Pred. No. 4.45e-04;
Matches 22; Conservative 60; Mismatches 56; Indels 1; Gaps 1;

Db 810 WYRGWGVDYGGGYNYNGKGRVTMADTSSNSRSVTAADTAVYVCVRGRSYDSGGDY 869
QY 256 tcaggggatcgatgcacagactctccagtatggcaataatcgatcccgattccgcgc 315
Db 870 WGGTAVTVSSHUVKDMTSSSSASVGDRTVTCRSSSTTHGNGNTYYWYKRAKRVSNRSGV 929
QY 316 tctcatctcc-atgttcgaccttcgttctctatctacgtcgctggaaggctgtgagcag 374

Db 930 SRSGSGSGTDYTTSSDATY 948
QY 375 gacgcagagaatcgcgttt 393

RESULT 15
ID PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DE
DT
Sequence 94, Application PC/TUS9511934
CC Sequence 94, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abitides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036

```

MILEAGE

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Mon Mar 8 21:44:20 1999; MasPar time 628.63 Seconds
1336.036 Million cell updates/sec

Tabular output not generated.

Title: >US-09-206-040-1
Description: (1-469) from US09206040A.seq
Perfect Score: 465
N.A. Sequence: 1 ttaactgcagcgcgcaggt.....gatgatacatggaagatnnt 469
Comp: aattgaacgtcgcnggtcca.....ctactatgtaactctctanna

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-est56
1:em_est1 2:em_gss1 3:em_gss2 4:em_gss3
genbank-est109
Database: 5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est20 17:gb_est21
18:gb_est22 19:gb_est23 20:gb_est24 21:gb_est25 22:gb_est26
23:gb_est27 24:gb_est28 25:gb_est29 26:gb_gss1 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 10.263; Variance 1.889; scale 5.433

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description Pred. No.
1	81	17.4	62 9 C74877 Rice cDNA, partial seq 4.58e-113
2	62	13.3	252 12 AA754459 97SN1787 Rice Immature 1.43e-74
3	52	11.2	252 12 AA754459 97SN1787 Rice Immature 3.86e-55
4	48	10.3	247 12 AA754458 97SN1784 Rice Immature 1.35e-47
5	45	9.7	247 12 AA754458 97SN1784 Rice Immature 4.80e-42
6	40	8.6	468 21 T18678 5C04C06-t7 membrane-fr 4.95e-33
7	40	8.6	568 21 T18698 5C04F08-t7 membrane-fr 4.95e-33
8	39	8.4	436 21 T18686 Human clone xs366 mRNA 2.86e-31
9	39	8.4	436 21 T18686 Human clone xs366 mRNA 2.86e-31
10	39	8.4	583 21 T14703 05C0402-t7 membrane-f 2.86e-31
11	39	8.4	589 9 H5299396 Homo sapiens mRNA; exp 2.86e-31
12	39	8.4	669 21 T18686 5C02H01-t7 membrane-fr 2.86e-31
13	38	8.2	163 21 T18683 5C04E02-t7 membrane-fr 1.59e-29

14	38	8.2	350 23 HSU46350 Human clone xs366 mRNA 1.59e-29
15	38	8.2	354 21 T14653 05C03g08-f21 membrane- 1.59e-29
16	38	8.2	370 23 HSU46347 Human clone xs361 mRNA 1.59e-29
17	38	8.2	373 21 T18695 5C04F05-t7 membrane-fr 1.59e-29
18	38	8.2	376 21 T18677 5C04B01-t7 membrane-fr 1.59e-29
19	38	8.2	379 21 T14744 05C04h07-f21 membrane- 1.59e-29
20	38	8.2	392 21 T18721 5C04H03-t7 membrane-fr 1.59e-29
21	38	8.2	400 23 HSU46349 Human clone xs363 mRNA 1.59e-29
22	38	8.2	404 21 T18665 5C02G02-t7 membrane-fr 1.59e-29
23	38	8.2	414 21 T18701 5C04G01-t7 membrane-fr 1.59e-29
24	38	8.2	414 21 T18700 5C04F10-t7 membrane-fr 1.59e-29
25	38	8.2	416 21 T18706 5C04G04-t7 membrane-fr 1.59e-29
26	38	8.2	430 23 HSU46341 Human clone xs353 mRNA 1.59e-29
27	38	8.2	431 21 T18292 5C06C07-t7 membrane-fr 1.59e-29
28	38	8.2	432 21 T18675 5C04A03-t7 membrane-fr 1.59e-29
29	38	8.2	432 21 T18703 5C04G02-t7 membrane-fr 1.59e-29
30	38	8.2	432 21 T18723 5C04H05-t7 membrane-fr 1.59e-29
31	38	8.2	436 21 T18699 5C04F09-t7 membrane-fr 1.59e-29
32	38	8.2	448 21 T18709 5C04G07-t7 membrane-fr 1.59e-29
33	38	8.2	450 21 T18295 5C06C11-t7 membrane-fr 1.59e-29
34	38	8.2	450 21 T18285 5C06B04-t7 membrane-fr 1.59e-29
35	38	8.2	457 21 T18704 5C04G03-t7 membrane-fr 1.59e-29
36	38	8.2	458 21 T18685 5C04E08-t7 membrane-fr 1.59e-29
37	38	8.2	458 21 T18684 5C04E07-t7 membrane-fr 1.59e-29
38	38	8.2	485 21 T18726 5C04F09-t7 membrane-fr 1.59e-29
39	38	8.2	550 21 T18696 5C04H06-t7 membrane-fr 1.59e-29
40	38	8.2	551 23 AA220909 24 Pinus radiata somat 1.59e-29
41	38	8.2	579 18 AI142137 GDM00L51 human embryo 1.59e-29
42	38	8.2	744 23 AA220873 13 Pinus radiata post 1.59e-29
43	38	8.2	815 18 AI142131 GDM00L19 human embryo 1.59e-29
44	38	8.2	881 18 AI142132 GDM00L03 human embryo 1.59e-29
45	38	8.2	895 18 AI142134 GDM00L05 human embryo 1.59e-29

ALIGNMENTS

RESULT 1 C74877 462 bp mRNA EST 18-SEP-1997
LOCUS Rice cDNA, partial sequence (E60595_1A), mRNA sequence.
DEFINITION C74877
ACCESSION
NID 92443106
KEYWORDS EST; EST(expressed sequence tag).
SOURCE Oryza sativa (strain:Nipponbare) Immature leaf including apical meristem (under short day condition) cDNA to mRNA.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; vascular plants; seed plants; Magnoliophyta;
Liliopsida; Poales; Poaceae; Oryza.
REFERENCE 1 (sites)
AUTHORS Sasaki,T. and Yamamoto,K.
TITLE Rice cDNA from Immature leaf including apical meristem
JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 462)
AUTHORS Sasaki,T.
TITLE Direct Submission
JOURNAL Submitted (13-AUG-1997) to the DDBJ/EMBL/GenBank databases. Takuji Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305, Japan (E-mail:tsasaki@affrc.go.jp, Tel:0298-38-7441, Fax:0298-38-7468)
PROJECT "RGP".
COMMENT Location/Qualifiers
FEATURES
source
1. 462
/organism="Oryza sativa"
/strain="Nipponbare"
/db_xref="taxon:4530"
/dev_stage="immature leaf including apical meristem (under short day condition)"
BASE COUNT 117 a 103 c 145 g 93 t 4 others
ORIGIN
Query Match 17.4%; Score 81; DB 9; Length 462;
Best Local Similarity 72.9%; Pred. No. 4.58e-113;

```

Matches 124; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
Db 50 TCGGANTGTCGCATGCTCGCATGTCGCCACATGGCTCTCTTACGTGCGC 109
Qy 300 tccgcgattcccgctcctcatctccatcttcgacatctcgcttctatcagtcgct 359
Db 110 GGCGCTGTGTCGCAGGATGCGCAGAGTAGGCTTTACCTGATTAAAGAGCTGCACAGCGGA 169
Qy 360 ggaaggctgtggcagcagcagaatcgctttatctcatcaagaagctgatagatc 419
Db 170 ACCGCCAGGCCATTCGACGATATCGGTGGATGATACCTTGAGAGTTGT 219
Qy 420 actgcccaggacaatctgctatctctggtgatgatacatgaagatnnt 469

RESULT 2
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION AA754459
NID 92801165
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
REFERENCE Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
AUTHORS Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 252)
Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
Unpublished (1998)
TITLE
JOURNAL
COMMENT
Contact: Eun M.Y.
Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.asti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhnahm@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
FEATURES
Location/Qualifiers
Source
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/notes="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional cDNA library inserted into Lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
BASE COUNT 5 a 21 c 12 g 35 t 179 others
ORIGIN
Query Match 13.3%; Score 62; DB 12; Length 252;
Best Local Similarity 9.9%; Pred. No. 1.43e-74;
Matches 22; Conservative 115; Mismatches 85; Indels 1; Gaps 1;
Db 29 BWVVCVASHGNVMSVHNTBRTGTHCDCKNVNWSNTWTGTVNWNVSGDWHVBNVNTKVD 88
Qy 215 cgttcataagattgcttgatttcattcaaaaatcgatcagggtgagtcacag 274
Db 89 VGNHTRCSRWRBVTMAHYHDYTNCCBBYNNNDYHMWHBMYBBTGCTMCWCBBHYNTK 148
Qy 275 actctccagtagggcaataaactcgatcccgctatcccgctcctcatcctcaggtgcg 334
Db 149 CTASGWHSTSTNYDKVKSNTW-GVTBSYDKSMHGWCBSBBVKHYHTKVSTTRATRSYTCVR 207

```

```

Qy 335 cacttcgctctatctacgtcgtaagcgctgtggcaggacgcagagaatcgcggtta 394
Db 208 KYCVMMTKKVKYKHVVBGCHBTDSKCKTMMTKNKHVMTST 250
Qy 395 tctcatcaagagctcgatgagtcacatcgctgcccaggagacaatct 437

RESULT 3
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION AA754459
NID 92801165
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
REFERENCE Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
AUTHORS Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 252)
Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
Unpublished (1998)
TITLE
JOURNAL
COMMENT
Contact: Eun M.Y.
Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.asti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhnahm@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
FEATURES
Location/Qualifiers
Source
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/notes="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional cDNA library inserted into Lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
BASE COUNT 5 a 21 c 12 g 35 t 179 others
ORIGIN
Query Match 11.2%; Score 52; DB 12; Length 252;
Best Local Similarity 9.5%; Pred. No. 3.86e-55;
Matches 19; Conservative 105; Mismatches 74; Indels 2; Gaps 2;
Db 41 MSVHNCBTRGTHCDCKNVNWSNTWTGTVNWNVSGDWHVBNVNTKVDVGNHTRCSRWRB 100
Cp 319 aggagcggggaatcgaggatcgatttgcctactgagatgctgctgactcgatcc 261
Db 101 VTRMAHYDYNCCBBYNNNDYHMWHBMYBBTGCTMCWCBBHYNTKCTASGWHSTNY 160
Cp 260 cctgactgctgatttgaatcaatgcaaatgcaatcttatgaacgat-tatgatattgt 202
Db 161 DVKSSTNTWGTBSYDKSMHGWCBSBBVKHYTKVSTTRATRSYTCVRKYCVMMTKKVK 220
Cp 201 atagttagtcagagagaattggtgattgattcgatcgatgataagcaagcaagcttgggt 142
Db 221 KYHVVBGCHBTDSKCKTMW 240
Cp 141 gctgcggttgcctactttct 122

```

[illegible]

REFERENCE 1 (bases 1 to 468)
 AUTHORS Shen,B., Carneiro,N., Torres-Jerez,I., Stevenson,R., McCreery,T., Helentjaris,T., Baysdorfer,C., Almira,E., Ferl,R., Habben,J. and Larkins,B.
 TITLE Partial sequencing and mapping of clones from two maize cDNA libraries
 JOURNAL Plant Mol. Biol. 26, 1085-1101 (1994)
 MEDLINE 95111093
 COMMENT Contact: The Maize cDNA Project

Helentjaris TG (primary contact)
 Dept. of Plant Sciences
 University of Arizona
 Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721
 ph: 602-6218-746
 fax: 602-621-7186
 E-mail: helnjars@ccit.arizona.edu

Chris Baysdorfer
 Department of Biological Sciences, School of Science
 California State University, Hayward
 Hayward, CA 94542
 ph: 510-881-3459
 fax: 510-727-2035
 E-mail: cbaysdor@sl.csu Hayward.edu

Rob Ferl
 Interdisciplinary Center for Biotechnology Research
 DNA Sequencing Core
 University of Florida
 P.O. Box 100695
 Gainesville, FL 32611-0695
 ph: 904-392-1928, ext. 301
 fax: 904-392-4072
 E-mail: robferl@nervm.nerdc.ufl.edu

Seq primer: T7.
 Location/Qualifiers
 1. 468
 /organism="Zea mays"
 /strain="W64A2"
 /note="Vector: ZipLox; Site_1: SalI; Site_2: NotI; ds-cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo- dT oligomer and then adding the second strand to RNase-nicked DNA:RNA hybrid with DNA Pol. SalI adaptors were added to the ends, the ds-cDNAs were then digested with NotI and size-selected. These were directionally-cloned into the ZipLox phage vector, excised as plasmids, and then individually analyzed."
 /db_xref="taxon:4577"
 /clone="5C04C06"
 /clone_lib="membrane-free polysomes from endosperm"
 /lab_host="DH10B"

BASE COUNT 89 a 110 c 123 g 99 t 47 others
 Query Match 8.6%; Score 40; DB 21; Length 468;
 Best Local Similarity 89.6%; Pred. No. 4.95e-33;
 Matches 43; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 10 CAGGTACGGTCCGGAATCCCGGTGACCCACGCGTCCGAGGCGT 57
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Qy 16 caggtancggtcaggaaattcccggtgacccacgctccgtagcgct 63

RESULT 7
 LOCUS T18698 568 bp mRNA EST 17-OCT-1996
 DEFINITION 5C04F08-T7 membrane-free polysomes from endosperm Zea mays cDNA clone 5C04F08 5' end similar to 22kd alpha zein alpha precursor,

ACCESSION T18698
 NID 9485628
 KEYWORDS EST.
 SOURCE maize.
 ORGANISM Zea mays
 Eukaryotae; mitochondrial eukaryotes; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
 REFERENCE 1 (bases 1 to 568)
 AUTHORS Shen,B., Carneiro,N., Torres-Jerez,I., Stevenson,R., McCreery,T., Helentjaris,T., Baysdorfer,C., Almira,E., Ferl,R., Habben,J. and Larkins,B.
 TITLE Partial sequencing and mapping of clones from two maize cDNA libraries
 JOURNAL Plant Mol. Biol. 26, 1085-1101 (1994)
 MEDLINE 95111093
 COMMENT Contact: The Maize cDNA Project

Contact: The Maize cDNA Project

Helentjaris TG (primary contact)
 Dept. of Plant Sciences
 University of Arizona
 Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721
 ph: 602-6218-746
 fax: 602-621-7186
 E-mail: helnjars@ccit.arizona.edu

Chris Baysdorfer
 Department of Biological Sciences, School of Science
 California State University, Hayward
 Hayward, CA 94542
 ph: 510-881-3459
 fax: 510-727-2035
 E-mail: cbaysdor@sl.csu Hayward.edu

Rob Ferl
 Interdisciplinary Center for Biotechnology Research
 DNA Sequencing Core
 University of Florida
 P.O. Box 100695
 Gainesville, FL 32611-0695
 ph: 904-392-1928, ext. 301
 fax: 904-392-4072
 E-mail: robferl@nervm.nerdc.ufl.edu

Rob Ferl
 Interdisciplinary Center for Biotechnology Research
 DNA Sequencing Core
 University of Florida
 P.O. Box 100695
 Gainesville, FL 32611-0695
 ph: 904-392-1928, ext. 301
 fax: 904-392-4072
 E-mail: robferl@nervm.nerdc.ufl.edu

Chris Baysdorfer
 Department of Biological Sciences, School of Science
 California State University, Hayward
 Hayward, CA 94542
 ph: 510-881-3459
 fax: 510-727-2035
 E-mail: cbaysdor@sl.csu Hayward.edu

Seq primer: T7.
 Location/Qualifiers
 1. 568
 /organism="Zea mays"
 /strain="W64A2"
 /note="Vector: ZipLox; Site_1: SalI; Site_2: NotI; ds-cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo- dT oligomer and then adding the second strand to RNase-nicked DNA:RNA hybrid with DNA Pol. SalI adaptors were added to the ends, the ds-cDNAs were then digested with NotI and size-selected. These were directionally-cloned into the ZipLox phage vector, excised as plasmids, and then individually analyzed."
 /db_xref="taxon:4577"
 /clone="5C04F08"
 /clone_lib="membrane-free polysomes from endosperm"
 /lab_host="DH10B"

BASE COUNT 153 a 183 c 104 g 123 t 5 others
 Query Match 8.6%; Score 40; DB 21; Length 568;
 Best Local Similarity 95.3%; Pred. No. 4.95e-33;
 Matches 41; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 7
 LOCUS T18698 568 bp mRNA EST 17-OCT-1996
 DEFINITION 5C04F08-T7 membrane-free polysomes from endosperm Zea mays cDNA clone 5C04F08 5' end similar to 22kd alpha zein alpha precursor,

Helentjaris TG (primary contact)
Dept. of Plant Sciences

Shen, B., Carneiro, N., Torres-Jerez, I., Stevenson, R., McCreery, T., Telent-Jaris, T., Baysdorfer, C., Almira, E., Perl, R., Habben, J. and Marking, B.

Gainesville, FL 32611-0695
ph: 904-392-1928, ext. 301
fax: 904-392-4072
E-mail: robferlner@nervm.nerdc

Seq primer: T7.

```

FEATURES
source
Location/Qualifiers
1..669
/organism="Zea mays"
/strain="W64A2"
/note="Vector: ZipLox; Site_1: SalI; Site_2: NotI; ds-cDNA
was prepared from oligo-dT selected mRNA by priming with a
NotI oligo- dT oligomer and then adding the second strand
to RNase nicked DNA:RNA hybrid with DNA PolI. SalI
adaptors were added to the ends, the ds-cDNAs were then
digested with NotI and size-selected. These were
directionally-cloned into the ZipLox phage vector, excised
as plasmids, and then individually analyzed."

```

```
/clone_lib="membrane-free polysomes from endospore"
```

BASE COUNT	ORIGIN	125 a	118 c	146 g	153 t	127 others
mrna	<1. .>669					

Query Match 8.4%; Score 39; DB 21; Length 669;
Best Local Similarity 75.8%; Pred. No. 2.86e-31;
Matches 47; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Db	12	CAGGTACCGGTCGGAATTC	CGGGTCGACCC	ACGGTCCGGATGC	AAGNANGNC	NAN	71	
Qy	16	caggtan	cggtcaga	attcccg	gtccac	cgcgtc	cgatcgacgac	75

Db 72 AG 73
76 ag 77

RESULT	13	EST	17-OCT-1996
LOCUS	TL18683	163 bp mRNA	
DEFINITION	5C04E02-T7 membrane-free polysomes from endosperm Zea mays cDNA clone 5C04E02 5' end, mRNA sequence.		

ACCESSION	TI8683
NID	g485613
KEYWORDS	EST.
SOURCE	maize.

ORGANISM	zeal mayas
REFERENCE	Eukaryotes: mitochondrial eukaryotes: Viridiplantae: Charophyta/Embryophyta group; Embryophyta: Magnoliophyta; Liliopsida; Poales; Poaceae; Zea. 1 (bases 1 to 163) Shen, B., Carneiro, N., Torres-Jerez, I., Stevenson, R., McCreery, T., Helentjaris, T., Baysdorfer, C., Almira, E., Perl, R., Habben, J., and Larkin, B.

TITLE	Partial sequencing and mapping of clones from two maize cDNA libraries
JOURNAL	Plant Mol. Biol. 26, 1085-1101 (1994)
MEDLINE	95111093

Contact: The Maize cDNA Project

Helentjaris TG (primary contact)
Dept. Of Plant Sciences
University of Arizona
Dept. Of Plant Sciences, University of Arizona, Tucson, AZ, 85721
ph: 602-6218-7446
fax: 602-621-7186
E-mail: helnjars@cicit.arizona.edu

Chris Baysdorfer

Department of Biological Sciences, School of Science
California State University, Hayward
Hayward, CA 94542
ph: 510-881-3459
fax: 510-727-2035
E-mail: chaysdor@es1.csu Hayward.edu

Rob Ferl
Interdisciplinary Center for Biote
DNA Sequencing Core
University of Florida
P.O. Box 100695
Gainesville, FL 32611-0695
ph: 904-392-1928, ext. 301
fax: 904-392-4072
E-mail: roberflernvm.nerdc.ufl.edu

Seq primer: T7.

```

FEATURES
    source
        Location/Qualifiers
            1..163
                /organism="Zea mays"
                /strain="W64A2"
                /note="Vector: ziplox; Site_1: SalI; Site_2: NotI; ds-cDNA
was prepared from oligo-dT selected mRNA by priming with a
NotI oligo- dT oligomer and then adding the second strand
to RNase-nicked DNA:RNA hybrid with DNA POLI. SalI
adaptors were added to the ends, the ds-cDNAs were then
digested with NotI and size-selected. These were
directionally-cloned into the ziplox phage vector, excised
as plasmids, and then individually analyzed."
                /db_xref="taxon:4577"
                /clone="5C04E02"
                /clone_lib="membrane-free polysomes from endosperm"
                /lab_host="DH10A"

```

BASE COUNT	ORIGIN	27 g	16 t	5 others
92 a	23 c	27 g	16 t	5 others

Query Match 8.2%; Score 38; DB 21; Length 163;
Best Local Similarity 95.1%; Pred. No. 1.59e-29;
Matches 39; Conservative 0; Mismatches 2; Indels

Db 11 CAGGTACCGGTCCGGAATTCGGGGTCGACCCACGCGTCCG 5'

Qy 16 caggtancggtcaggaattcccggttcgacccacgctccg 5'

RESULT 1.

ACCESSION	1*	HSU46350	360 bp	mrna	EST	13-MAR-1997
LOCUS		Human Clone	xs366	mrna sequence.		
DEFINITION		U46350				
ACCESSION		U46350				
NID		q1236492				
KEYWORDS		EST.				
SOURCE		human.				
ORGANISM		Homo sapiens				
		Eukaryotae:	mitochondrial eukaryotes:	Metazoa:	Chordata:	
		Vertebrata:	Eutheria:	Primates:	Catarrhini:	Hominidae: Homo.
REFERENCE		1 (bases 1 to 360)				
AUTHORS		Gress,T.M., Muller-Pillasch,F., Geng,M., Zimmerhackl,F., Zehner,G., Friess,H., Buchler,M., Adler,G. and Lehrach,H.				
TITLE		A pancreatic cancer-specific expression profile				
JOURNAL		Oncogene 13 (8), 1819-1830 (1996)				

FEATURES
source

```

/organism="Homo sapiens"
/note="overexpressed in pancreatic cancer"

```

```

/db_xref="taxon:9606"
/cell_line="Patu 8988t"
/tissue_type="pancreatic cancer"
/clone="xs366"

BASE COUNT      53 a      95 c      129 g      67 t      16 others
ORIGIN

Query Match      8.2%; Score 38; DB 23; Length 360;
Best Local Similarity 95.1%; Pred. No. 1.59e-29;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 27 CAGGTACCGTCCGGAATTCGCCGGTCCGACCCACGCGTCCG 67
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 16 caggtancggtcagggaattcccggtcgaccacgcgtccg 56

RESULT 15
LOCUS T14653 364 bp mRNA EST 17-OCT-1996
DEFINITION 05c03g08-f21 membrane-free polysomes from endosperm Zea mays cDNA
clone 05c03g08 3' end similar to starch branching enzyme II, mRNA
sequence.
ACCESSION T14653
NID 9440632
KEYWORDS EST.
SOURCE maize.
ORGANISM Zea mays
Eukaryotes; mitochondrial eukaryotes; Viridiplantae;
Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
Liliopsida; Poales; Poaceae; Zea.
REFERENCE 1 (bases 1 to 364)
AUTHORS Shen, B., Carneiro, N., Torres-Verez, I., Stevenson, R., McCreery, T.,
Helentjaris, T., Baysdorfer, C., Almira, E., Perl, R., Habben, J. and
Larkins, B.
TITLE Partial sequencing and mapping of clones from two maize cDNA
libraries
JOURNAL Plant Mol. Biol. 26, 1085-1101 (1994)
MEDLINE 95111093
COMMENT Other ESTs: 05c03g08-t7
Contact: The Maize cDNA Project

```

Search completed: Mon Mar 8 22:09:57 1999
Job time : 1537 secs.

```

mRNA
BASE COUNT      94 a      86 c      86 g      93 t      5 others
ORIGIN

Query Match      8.2%; Score 38; DB 21; Length 364;
Best Local Similarity 95.1%; Pred. No. 1.59e-29;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 253 CGGACGCGTGGTCCGACCCGGGAATTCGGACCGGTACCTG 293
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 56 cggacgcgtgggtcgaccccggaattcctgaccgntacctg 16

Search completed: Mon Mar 8 22:09:57 1999
Job time : 1537 secs.

```

```

Chris Baysdorfer
Department of Biological Sciences, School of Science
California State University, Hayward
Hayward, CA 94542
ph: 510-881-3459
fax: 510-727-2035
E-mail: cbaysdor@sl.csbayward.edu

Rob Perl
Interdisciplinary Center for Biotechnology Research
DNA Sequencing Core
University of Florida
P.O. Box 100695
Gainesville, FL 32611-0695
ph: 904-392-1928, ext. 301
fax: 904-392-4072
E-mail: robperl@nerdc.ufl.edu
Seq primer: F21.
Location/Qualifiers
1..364
/organism="Zea mays"
/strain="W64A2"
/note="Vector: Ziplox; Site_1: SalI; Site_2: NotI; ds-cDNA
was prepared from oligo-dT selected mRNA by priming with a

```

 W I S E N E H

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Mon Mar 8 22:10:16 1999; MasPar time 74.90 Seconds
 Tabular output not generated. 851.874 Million cell updates/sec

Title: >US-09-206-040-1
 Description: (1-469) from US09206040A.seq
 Perfect Score: 465
 N.A. Sequence: 1 ttaacttcagcgcgncagggt.....gatgatacattgaagatnnt 469
 Comp: aattgaacgcgcgngtcca.....ctactatgtaactctanna

Scoring table: TABLE default
 Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 58026449 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: n-geneseq32
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39 40:part40

Statistics: Mean 8.266; Variance 4.624; scale 1.788

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
C 1	39	8.4	204	1	N81164	1.27e-10
C 2	38	8.2	91	9	O51746	5.35e-10
C 3	38	8.2	91	9	O51746	5.35e-10
C 4	38	8.2	570	37	T91275	5.35e-10
C 5	38	8.2	969	38	T96721	5.35e-10
C 6	38	8.2	1581	22	T32233	5.35e-10
C 7	38	8.2	1698	40	V17199	5.35e-10
C 8	34	7.3	204	1	N81164	1.52e-07
C 9	32	6.9	114	12	O70468	2.39e-06
C 10	30	6.5	114	12	O70468	3.58e-05
C 11	30	6.5	114	12	O70467	3.58e-05
C 12	30	6.5	114	12	O70467	3.58e-05
C 13	29	6.2	114	12	O70468	1.36e-04

c	14	29	6.2	114	12	Q70465	Generic DNA sequence	1.36e-04
c	15	28	6.0	114	12	Q70469	Generic DNA sequence <td>5.06e-04</td>	5.06e-04
c	16	28	6.0	114	12	Q70466	Generic DNA sequence <td>5.06e-04</td>	5.06e-04
c	17	28	6.0	114	12	Q70469	Generic DNA sequence <td>5.06e-04</td>	5.06e-04
c	18	28	6.0	114	12	T76405	Human endothelin-1 an	5.06e-04
c	19	27	5.8	114	12	Q70470	Generic DNA sequence <td>1.86e-03</td>	1.86e-03
c	20	27	5.8	114	12	Q70470	Generic DNA sequence <td>1.86e-03</td>	1.86e-03
c	21	27	5.8	2606	32	T72895	Plasmodium E3la gene	1.86e-03
c	22	27	5.8	2606	14	Q83527	P. falciparum E3la ge	1.86e-03
c	23	27	5.8	5030	36	T75702	pGreenlateral vector.	1.86e-03
c	24	26	5.6	114	12	Q70466	Generic DNA sequence	6.59e-03
c	25	26	5.6	188	32	T76270	Human MDNCF antisense	6.59e-03
c	26	26	5.6	190	32	T76452	Chymase antisense oli	6.59e-03
c	27	25	5.4	39	7	O51787	Mixed oligonucleotide	2.36e-02
c	28	25	5.4	114	12	Q70471	Generic DNA sequence	2.36e-02
c	29	25	5.4	114	12	Q70472	Generic DNA sequence	2.36e-02
c	30	25	5.4	114	12	Q70472	Generic DNA sequence	2.36e-02
c	31	25	5.4	114	12	T76405	Human endothelin-1 an	2.36e-02
c	32	25	5.4	565	6	Q35072	HCV envelope region n	2.36e-02
c	33	24	5.2	74	21	T13613	DC43 TSAR library gen	8.19e-02
c	34	24	5.2	81	21	T13611	DC43 TSAR library gen	8.19e-02
c	35	24	5.2	114	12	Q70473	Generic DNA sequence	8.19e-02
c	36	24	5.2	114	12	Q70473	Generic DNA sequence	8.19e-02
c	37	24	5.2	114	12	Q70471	Generic DNA sequence	8.19e-02
c	38	23	4.9	36	2	Q11195	Ballast Constituent c	2.78e-01
c	39	23	4.9	75	21	T13612	DC43 TSAR library gen	2.78e-01
c	40	23	4.9	82	21	T13610	DC43 TSAR library gen	2.78e-01
c	41	23	4.9	130	32	T76152	Human vascular cell a	2.78e-01
c	42	23	4.9	130	32	T76452	Chymase antisense oli	2.78e-01
c	43	22	4.7	75	21	T13612	DC43 TSAR library gen	9.18e-01
c	44	22	4.7	81	21	T13611	DC43 TSAR library gen	9.18e-01
c	45	22	4.7	89	32	T76219	Human IL5 antisense o	9.18e-01

ALIGNMENTS

RESULT 1
 ID N81164 standard; DNA; 204 BP.
 AC N81164;
 DT 08-NOV-1990 (first entry)
 DE Base substituted E.coli beta-galactosidase alpha-fragment.
 KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
 OS Escherichia coli.
 FH Key Location/Qualifiers
 FT misc_feature 19..69
 FT primer_bind 187..204
 FT /function= multiple cloning site
 FT /tag= a
 FT /tag= b
 EP-285123-A.
 PD 05-MAY-1988.
 PF 30-MAR-1988; 105163.
 PR 03-APR-1987; US-034819.
 PA (SUSO) SUOMEN SOKERI OY.
 PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
 DR WPI: 88-279927/40.
 PT Introducing random point mutations into nucleic acids -
 PT by prep of single stranded template, annealing a primer, elongation,
 PT misincorporation, completion of molecules and screening.
 PS Disclosure; P; English.
 CC Random point mutations were introduced into the alpha fragment of
 CC E.coli beta-galactosidase. The wild type sequence was obtained as a
 CC single stranded template and an oligonucleotide was hybridised to
 CC it to generate a popn of DNA molecules which terminate at all
 CC possible nucleotide positions within a specified region. The
 CC variable 3' ends generated in this way are used as primers for
 CC reverse transcriptase. Nucleotides are misincorporated by the
 CC transcriptase and the molecules are completed to forms that can be
 CC amplified and then expressed in a suitable host-vector system.
 CC The sequence covers all 176 diff base substitutions, most of which
 CC occurred singularly in any given mutant.
 CC See also P80575.
 SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Seq	Sequence	91 bp	3	17 C	13 G	4 T	1
	Query Match	8.2%	Score 38;	DB 9;	Length 91;		
	Best Local Similarity	11.7%;	Pred. No. 5.35e-10;				
	Matches	7;	Conservative 42;	Mismatches 11;	Indels 0;	Gaps 0;	
Db	2 gctccgcgcgshsvyvvhvvshhshsvhvhvvhvvvvhvvhvvhvvhvhyvsvvc	61					
	:						
Cp	141 gctgcgctgctactttcttctgccaagaagtcaccagaatccactctcttcccccc	82					

AC	
AD	131275;
AE	23-APR-1998 (first entry)
AF	cDNA encoding a novel beta-chemokine designated PTEC.
AG	Beta-chemokine; PTEC; homologue; human RANTES; C-C chemokine;
AH	diagnosis; treatment; activated T cell;
AI	immediate type hypersensitivity; leukocyte proliferation; ss.
AK	Homo sapiens.
AL	OS
AM	FH Key
AN	Location/Qualifiers
AO	110..391
AP	CDS
AQ	/*tag= a
AR	WO9739126-A1.
AS	PN
AT	23-OCT-1997.
AV	PD
AW	Pf 14-APR-1997; U06249.
AX	PR 17-APR-1996; US-633682.
AY	(INCY-) INCYTE PHARM INC.
AZ	PA Bandman O, Hawkins PR, Murry LE;
BA	PI WPI; 97-526462/48.
BB	DR P-SDBS; W21271.
BC	DR PTEC, a novel human beta-chemokine - useful for diagnosing and
BD	treating e.g. viral, bacterial, fungal infections, auto-immune
BE	diseases, etc
BF	PT PT
BG	PS Claim 2. Pages 39-40; 59pp; English.
BH	CC The present DNA sequence encodes a novel beta-chemokine polypeptide
BI	CC designated PTEC. The protein has Cys residues of a beta-chemokine (C32,
BJ	CC C33, C56 and C72). The present sequence was isolated from a cDNA library
BK	CC generated from non-tumorous human prostate tissue. The PTEC protein is
BL	CC

a homologue of human RANTES, which is a C-C chemokine with structural similarity to interleukin-8 and human MIP-1beta. The polynucleotide sequence encoding PTFC can be used to diagnose or treat a condition associated with the presence of activated T cells including viral, bacterial, fungal or helminthic infections, allergic or asthmatic responses, mechanical injury associated with trauma, arteriosclerosis, atherogenesis or collagen vascular diseases, autoimmune diseases such as rheumatoid arthritis, myasthenia gravis, systemic lupus erythematosus or haemolytic anaemia, leukaemia, lymphomas or carcinomas, and diseases of immediate type hypersensitivity which involve activation or excessive proliferation of leukocytes, particularly monocytes, macrophages, eosinophils, basophil, mast and T cells.

570 BP: 149 A: 155 C: 138 G: 128 T: 130

CC	Immediate type hypersensitivity which involve activation of excessive
CC	proliferation of leukocytes, particularly monocytes, macrophages,
CC	eosinophils, basophil, mast and T cells.
SC	Sequence 570 BP; 149 A; 155 C; 138 G; 128 T;
	Query Match 8.2%; Score 38; DB 37; Length 570;
	Best Local Similarity 95.1%; Pred. No. 5.35e-10;
	Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db	46 caggtaccggtccggattccgggtcgaccaccacggtccg 86
Qy	16 caggtancggtcagggaattccgggtcgaccaccacggtccg 56
RESULT	5
ID	T96721 standard; cDNA; 969 BP.
ID	T96721;
AC	AC
DT	27-APR-1998 (first entry)
DE	Human synaptogyrin cDNA.
KW	Synaptogyrin; human; colon cancer; lung cancer; brain tumour;
KW	Crohn's disease; rheumatoid arthritis; AIDS; allergy; urticaria;
KW	juvenile diabetes; scleroderma; antitense; diagnosis; therapy; ss.
OS	Homo sapiens.
FT	Key
FT	Location/Qualifiers
CDS	110..784

```

FT W09741143-A1. /*tag= a
PN 06-NOV-1997.
PD 30-APR-1997; U07378.
PF 30-APR-1996; US-700637.
PR (INCY-) INCYTE PHARM INC.
PA Hawkins PR, Murry LE, Stuart SG:
PI WPI: 97-549677/50.
DR P-PSDB; W36516.
PT Novel human synaptogyrin homologue - useful for diagnosis and
PT treatment of, e.g. rheumatoid arthritis, Crohn's disease, cancers,
PT etc.
PS Claim 2; Page 46-47; 74pp; English.
CC This polynucleotide (snpg) identifies and encodes a novel human
CC synaptogyrin homologue (SNPG) (see W36516). It was initially
CC identified in Incyte Clone 775426 from a colon cDNA library
CC (COLN0705) prepared from a Crohn's disease patient. Also claimed
CC are: (1) an expression vector containing PN; (2) a host cell
CC transformed with the vector of (1); (3) the purified 224 amino acid
CC SNPG protein; (4) an antisense molecule complementary to PN; (5) an
CC antibody (Ab) specific for the polypeptide of (3); and (6) an
CC inhibitor specific for the polypeptide of (3). Human SNPG can be
CC used for the diagnosis and treatment of a disease associated with
CC synaptogyrin, especially Crohn's disease, rheumatoid arthritis,
CC asthma and cancers or tumours of the lung, colon or brain (claimed).
CC The antisense molecule, Ab or inhibitor can be used to treat, e.g.
CC AIDS, allergy, urticaria, juvenile diabetes, scleroderma, etc.
SQ Sequence 969 BP; 163 A; 324 C; 280 G; 202 T;

Query Match 8.2%; Score 38; DB 38; Length 969;
Best Local Similarity 95.1%; Pred. No. 5.35e-10;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 46 caggtaccggtccggaattcccggtcgaccacgcgctccg 86
QY 16 caggtancggtcagggaattcccggtcgaccacgcgctccg 56

RESULT 6
ID T32233 standard; cDNA: 1581 BP.
AC T32233;
DT 26-OCT-1996 (first entry)
DE Cuphea C14:0-ACP thioesterase cDNA clone MCR34 (CpRatB2).
KW Myristate; myristic acid; fatty acid; acyl-ACP thioesterase;
KW seed oil; oilseed; vegetable oil; Brassica; transgenic plant;
KW surfactant; ss.
OS Cuphea palustris.
FH Key Location/Qualifiers
FT cds 119..1354
FT /*tag= a
FT W09623892-A2.
PN 08-AUG-1996.
PD 01-FEB-1996; U01585.
PF 02-FEB-1995; US-383756.
PR 05-JUN-1995; US-460898.
PA (CALJ ) CALGENE INC.
PI Dehesh K, Hawkins D, Voelker TA;
DR WPI: 96-371439/37.
DR P-PSDB; W02081.
PT Recombinant prodn. of myristate in plant cells - using DNA with
PT preferential activity on C14 fatty acids from Cuphea palustris,
PT nutmeg and camphor, useful in detergent and food industries
PS Example 1: Fig 1A-1E; 77pp; English.
CC A cDNA clone (T32233), designated MCR34 (pCpRatB2), codes for Cuphea
CC palustris C14:0-ACP thioesterase (W02081), an enzyme that acts
CC primarily on C14:0-ACP substrates, forming C14:0 (myristate). It
CC was isolated from a developing seed cDNA library by screening with
CC sequences encoding medium-chain acyl-ACP thioesterases from Cuphea
CC hookeriana. Constructs for expression of clone MCR34 in plant seeds
CC under the control of napin or oleosin regulatory regions were prepd.
CC These allow high-level prodn. of myristate (useful in surfactants
CC and foods) in plant cells, e.g. Brassica.
SQ Sequence 1581 BP; 418 A; 356 C; 410 G; 397 T;

```

```

Query Match 8.2%; Score 38; DB 22; Length 1581;
Best Local Similarity 95.1%; Pred. No. 5.35e-10;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 41 caggtaccggtccggaattcccggtcgaccacgcgctccg 81
QY 16 caggtancggtcagggaattcccggtcgaccacgcgctccg 56

RESULT 7
ID V17199 standard; cDNA: 1698 BP.
AC V17199;
DT 10-JUN-1998 (first entry)
DE Human MKK3-interacting protein (MIP) encoding cDNA.
KW MKK3-interacting protein; MIP; mitogen activated protein kinase;
KW MAPK kinase-3; human; transduction; target; extracellular signal;
KW interfering mutant; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 167..1213
FT /*tag= a
FT /product= "MKK3-interacting protein"
FT W09800339-A2.
PN 08-JAN-1998.
PD 02-JUL-1997; U10866.
PR 01-JUL-1997; US-886572.
PR 03-JUL-1996; US-021224.
PR 12-JUL-1996; US-021641.
PA (CHIR ) CHIRON CORP.
PI Kavanaugh MW, Shyamala V;
DR WPI: 98-086965/08.
DR P-PSDB; W52169.
PT Human MKK3-interacting protein, used to identify human tissues in
PT malignant conditions - comprises nucleic acid, antibodies, fusion
PT proteins and peptide fragments
PS Disclosure; Page 14; 17pp; English.
CC This cDNA encodes a human MKK3-interacting protein (MIP). The mitogen-
CC activated protein kinase (MAPK) kinase-3 (MKK3) is a protein kinase that
CC phosphorylates p38 MAP kinase specifically and is involved in transducing
CC stress signals. The MIP binds to and interacts with a dominant
CC interfering mutant form of MKK3 and may be involved in transducing
CC extracellular signals to the nucleus, resulting in activation of p38
CC kinase. MIP is expressed in human brain, kidney, liver, lung, pancreas
CC and spleen but not in heart or striated muscle. Detection of MIP in a
CC body sample (particularly a tumour or metastasis) is used to identify the
CC tissue source, the tissue selected from the group consisting of brain,
CC kidney, liver, lung, pancreas or spleen tissue. Fusion proteins comprising
CC at least 8 contiguous amino acids of MIP fused to a second protein
CC sequence can be used to target MIP, or its peptides, to specific cells or
CC tissues. They can be used as a source of MIP in assays and also as
CC immunogen. The MIP encoding nucleic acid is used for expression of
CC recombinant MIP polypeptides and as a source of probes for detecting
CC MIP-related mRNA.
SQ Sequence 1698 BP; 455 A; 381 C; 449 G; 413 T;

```

```

Query Match 8.2%; Score 38; DB 40; Length 1698;
Best Local Similarity 95.1%; Pred. No. 5.35e-10;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 62 caggtaccggtccggaattcccggtcgaccacgcgctccg 102
QY 16 caggtancggtcagggaattcccggtcgaccacgcgctccg 56

RESULT 8
ID N81164 standard; DNA: 204 BP.
AC N81164;
DT 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
FH Key Location/Qualifiers

```

```

FT misc_feature 19..69 /*tag= a
FT /function= multiple cloning site
FT primer_bind 187..204
FT /*tag= b
PD EP-285123-A.
PN 05-MAY-1988.
PF 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P., Knowles J., Koivula A., Bamford J., Reinikainen T.;
DR WPI; 88-279927/40.
PT Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure: p: English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80575
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 7.3%; Score 34; DB 1; Length 204;
Best Local Similarity 5.7%; Pred. No. 1.52e-07;
Matches 5; Conservative 48; Mismatches 35; Indels 0; Gaps 0;

Db 98 rbnyrdynrdaaawcyrrsvkydcynachddhlyvbybbvynhnnncocbn 157
: : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 8 gcagcgcaggtancggtcaggaatccggtgacccacgcgctcgtagcggtcgca 67
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 158 nhvchvbnhbnhwayrvhrdarddhv 185
: : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 68 aagacacagagaagggggggaagaga 95
: : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
ID Q70468 standard; DNA; 114 BP.
AC Q70468;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN W09418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UNYC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65154.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure: Page 35; 255pp; English.
CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)11Z(NNB)14(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70468-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with

```

```

CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)11(TGC)(NNB)6Z(NNB)7(TGC)(NNB)10Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compps. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 6.9%; Score 32; DB 12; Length 114;
Best Local Similarity 3.6%; Pred. No. 2.39e-06;
Matches 4; Conservative 32; Mismatches 76; Indels 0; Gaps 0;

Db 3 bnnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn 62
: : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 308 tccgcgcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 367
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 bnnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn 114
: : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 368 gtgcagacagacagacagacagacagacagacagacagacagacagacagac 419
: : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
ID Q70465 standard; DNA; 114 BP.
AC Q70465;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN W09418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UNYC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65150 and R65151.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure: Page 35; 255pp; English.
CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)11Z(NNB)14(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with

```

CC affinity for a ligand and a second effector peptide portion that is
 CC chemically or biologically active. They may further comprise a linker
 CC peptide between the 2 domains. The oligonucleotides are also designed so
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned
 CC in, or flanking, the unpredicted or variant residues. These residues
 CC confer some degree of conformational rigidity to the peptides. The TSARs
 CC or compns. comprising a TSAR binding domain can be used in vivo to
 CC deliver a chemically or biologically active moiety, eg. metal ion,
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
 CC cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need
 CC for complex methods of hybridoma formation or in vivo antibody
 CC production. The TSARs are easily characterised and have designed
 CC activity allowing direct and rapid detection in a screening process.
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 6.5%; Score 30; DB 12; Length 114;
 Best Local Similarity 3.6%; Pred. No. 3.58e-05;
 Matches 4; Conservative 31; Mismatches 77; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbntgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnn 62
 QY 287 ggcaataatcgatccgcattccgcgctccatctcattcgttcgcaattcgttc 346
 Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbntgcnnbnnbnnb 114
 QY 347 tatctacgtcgtggaaggctgtggcaggacgcagagaatcggtttatctc 398

RESULT 11

ID Q70467 standard; DNA; 114 BP.
 AC Q70467;
 DT 05-APR-1995 (first entry)
 DE Generic DNA sequence to generate a random TSAR peptide library.
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; generic; ss.
 OS Synthetic.
 FH Key
 FT misc_feature 55..60
 FT /*tag= a
 FT /note= "this sequence represents '2'; 2 can be a
 FT sequence of 6, 9 or 12 nucleotides (see
 FT comments)"

WO9418318-A.
 PD 18-AUG-1994.
 PF 01-FEB-1994; U00977.
 PR 01-FEB-1993; US-013416.
 PR 30-DEC-1993; US-176500.
 PR 31-JAN-1994; US-189331.
 PA (UYNC-) UNIV NORTH CAROLINA.
 PI Fowlkes DM, Kay BK;
 DR WPI: 94-279739/34.
 DR P-PSDB; R65153.
 PT Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PT comprising a binding domain and an effector domain
 PS Disclosure: Page 35: 255pp; English.
 CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be
 CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)1Y. X
 CC and Y are flanking restriction sites (X is not the same as Y) that are
 CC not specified further. Other generic sequences are shown in Q70466-68.
 CC Other specific peptides generated by these generic sequences are shown in
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
 CC comprising at least two functional regions - a binding domain with
 CC affinity for a ligand and a second effector peptide portion that is
 CC chemically or biologically active. They may further comprise a linker
 CC peptide between the 2 domains. The oligonucleotides are also designed so
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned
 CC in, or flanking, the unpredicted or variant residues. These residues
 CC confer some degree of conformational rigidity to the peptides. The TSARs
 CC or compns. comprising a TSAR binding domain can be used in vivo to

CC deliver a chemically or biologically active moiety, eg. metal ion,
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
 CC cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need for
 CC complex methods of hybridoma formation or in vivo antibody production.
 CC The TSARs are easily characterised and have designed activity allowing
 CC direct and rapid detection in a screening process.
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 6.5%; Score 30; DB 12; Length 114;
 Best Local Similarity 1.9%; Pred. No. 3.58e-05;
 Matches 2; Conservative 31; Mismatches 73; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbntgcnnbnnbnnnnnn 62
 QY 305 cattccgcgctccatctcattcgttcgcaattcgttcattcgttcgctggaag 364
 Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbntgcnnbnnbnnb 108
 QY 365 gctgtggcaggacgcagagaatcggtttatctcattcattcattcattcattc 410

RESULT 12

ID Q70467 standard; DNA; 114 BP.
 AC Q70467;
 DT 05-APR-1995 (first entry)
 DE Generic DNA sequence to generate a random TSAR peptide library.
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; generic; ss.
 OS Synthetic.
 FH Key
 FT misc_feature 55..60
 FT /*tag= a
 FT /note= "this sequence represents '2'; 2 can be a
 FT sequence of 6, 9 or 12 nucleotides (see
 FT comments)"

WO9418318-A.
 PD 18-AUG-1994.
 PF 01-FEB-1994; U00977.
 PR 01-FEB-1993; US-013416.
 PR 30-DEC-1993; US-176500.
 PR 31-JAN-1994; US-189331.
 PA (UYNC-) UNIV NORTH CAROLINA.
 PI Fowlkes DM, Kay BK;
 DR WPI: 94-279739/34.
 DR P-PSDB; R65153.
 PT Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PT comprising a binding domain and an effector domain
 PS Disclosure: Page 35: 255pp; English.
 CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be
 CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)1Y. X
 CC and Y are flanking restriction sites (X is not the same as Y) that are
 CC not specified further. Other generic sequences are shown in Q70466-68.
 CC Other specific peptides generated by these generic sequences are shown in
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
 CC comprising at least two functional regions - a binding domain with
 CC affinity for a ligand and a second effector peptide portion that is
 CC chemically or biologically active. They may further comprise a linker
 CC peptide between the 2 domains. The oligonucleotides are also designed so
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned
 CC in, or flanking, the unpredicted or variant residues. These residues
 CC confer some degree of conformational rigidity to the peptides. The TSARs
 CC or compns. comprising a TSAR binding domain can be used in vivo to
 CC deliver a chemically or biologically active moiety, eg. metal ion,
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
 CC cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need for
 CC complex methods of hybridoma formation or in vivo antibody production.
 CC The TSARs are easily characterised and have designed activity allowing
 CC direct and rapid detection in a screening process.

Cp 148 ctttggctggttgacttttcgtgcagaaagtccacaagaatcccactccttt 89
Db 66 bnnbnbnbnbnbtgnnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn 114
Cp 88 ccccccttctgtcgcttttcgcagccgtacgacgctggtgac 40

RESULT 14

ID	Q70465 standard; DNA; 114 BP.
AC	Q70465;
DT	O5-APR-1995 (first entry)
DE	Generic DNA sequence to generate a random TSAR peptide library.
KW	TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW	effector domain; concatenated heterofunctional protein; linker;
KS	direct; rapid; detection; screening; treatment; generic; ss.
OS	Synthetic.
Key	Location/Qualifiers
FH	misc_feature 55..60
FT	/tag= a
FT	/note= "this sequence represents 'z'; z can be a
FT	sequence of 6, 9 or 12 nucleotides (see
FT	comments)"

PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PR (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PDSB; R65150 and R65151.
DR Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 25Sep; English.
CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)V6(TGC)(NNB)112((NNB)J14(TGC)(NNB))3V. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown i
CC R65151-94. TSARS are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active.They may further comprise a linker
CC peptide between the 2 domains.The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compans. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARS are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 6.2%; Score 29; DB 12; Length 114;
Best Local Similarity 4.3%; Pred.No.1.36e-04;
Matches 5; Conservative 28; Mismatches 70; Indels 0; Gaps 0;

Db 12 bnnbnbnbtgnnbn 71
Cp 149 tctttgctgcggttgtacctttctgtgcagaaagtcaccaagaatcccactcctt 90
Db 72 bnnbn 114
Cp 89 tcctccccccttcctgtcgtcttttcgcagcgcttacgcacgcqtg 47

Search completed: Mon Mar 8 22:14:25 1999
Job time : 249 secs.

THIS PAGE BLANK (USPTO)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Mon Mar 8 22:40:59 1999; MasPar time 739.30 seconds
Tabular output not generated. 1499.157 Million cell updates/sec

Title: >US-09-206-040-1
Description: (1-469) from US09206040A.seq
Perfect Score: 465
N.A. Sequence: 1 ttaacttcgacgncacagt.....gatgatacattgaagatntt 469
Comp: aattgaacgctcgcngtcca.....ctactatgtaactttatanna

Scoring table: TABLE jmetric

Gap 60

Nmatch STD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl57

1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_om 8:em_ov 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_un 15:em_v1

Database: genbank110

16:gb_ba1 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_om 21:gb_ov
22:gb_pat 23:gb_ph 24:gb_pl1 25:gb_pl2 26:gb_pl1
27:gb_pr2 28:gb_pr3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
33:gb_un 34:gb_v1

Statistics: Mean 8.428; Variance 2.694; scale 3.128

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	30	6.5	2125	28	AB016247 Homo sapiens mRNA for	6.08e-11
2	28	6.0	367	31	HSRSTS367 H. sapiens STS, 367bp,	5.65e-09
3	28	6.0	372	27	HSXS16 H. sapiens (xs16) mRNA,	5.65e-09
4	28	6.0	770	27	HSU77594 Human tazarotene-induc	5.65e-09
5	28	6.0	1009	24	AB010946 Arabidopsis thaliana m	5.65e-09
6	28	6.0	1214	29	AF019249 Mus musculus Nmi mRNA,	5.65e-09
7	28	6.0	1372	25	AF034255 Arabidopsis thaliana r	5.65e-09
8	28	6.0	1382	24	AFRSP40 Arabidopsis thaliana m	5.65e-09
9	28	6.0	1433	25	AF029216 Arabidopsis thaliana p	5.65e-09
10	28	6.0	1581	22	I60018 Sequence 1 from patent	5.65e-09
11	28	6.0	1640	24	AB000453 Petunia hybrida mRNA f	5.65e-09
12	28	6.0	1709	29	AF031814 Mus musculus pregnane	5.65e-09
13	28	6.0	1779	29	AF071316 Mus musculus COP9 comp	5.65e-09

14	28	6.0	2111	25	AF071862	Vigna unguiculata phos	5.65e-09
15	28	6.0	2351	25	AF060518	Cuphea pulcherrima 3-k	5.65e-09
16	28	6.0	2879	27	AF004231	Homo sapiens monocyte/	5.65e-09
17	28	6.0	3410	26	HSU27699	Human pepBCT-1 betain	5.65e-09
18	28	6.0	8665	28	HSU86136	Human telomerase-assoc	5.65e-09
19	27	5.8	1291	25	ATRNABIND	Arabidopsis thaliana m	5.21e-08
20	27	5.8	1500	21	XLNAKATP	X. laevis mRNA for Na,K	5.21e-08
21	27	5.8	1669	29	MMCARHOM	M. musculus mRNA for co	5.21e-08
22	27	5.8	2344	20	OCU36790	Oryctolagus cuniculus	5.21e-08
23	26	5.6	1829	24	ATU72153	Arabidopsis thaliana b	4.64e-07
24	22	4.7	1836	24	ATU39783	Arabidopsis thaliana a	1.97e-03
25	21	4.5	1257	22	I89416	Sequence 11 from patent	1.42e-02
26	21	4.5	1498	21	DRSAPK3	Danio rerio mRNA for s	1.42e-02
27	21	4.5	1627	20	PIGHRHR	Porcine growth hormone	1.42e-02
28	21	4.5	1712	24	ATAJ2414	Arabidopsis thaliana m	1.42e-02
29	21	4.5	2055	22	I89410	Sequence 3 from patent	1.42e-02
30	21	4.5	2345	29	AF090334	Mus musculus FK506 bin	1.42e-02
31	20	4.3	4701	29	AF022992	Mus musculus Rigu1 MRN	9.70e-02
32	20	4.3	5565	21	GGCOLVIA	Chicken COLVIA2 gene f	9.70e-02
33	20	4.3	41220	27	AC004777	Homo sapiens chromosom	9.70e-02
34	20	4.3	9482	25	ATF26P21	Arabidopsis thaliana D	9.70e-02
35	19	4.1	270	26	HSCHBFC	H. sapiens mRNA for CCA	6.24e-01
36	19	4.1	588	21	XLAJ9291	Xenopus laevis cDNA cl	6.24e-01
37	19	4.1	1692	29	MMU87814	Mus musculus PEST phos	6.24e-01
38	19	4.1	1708	19	TRC223614	Tribolium castaneum MR	6.24e-01
39	19	4.1	2509	21	XLJ57455	Xenopus laevis growth	6.24e-01
40	19	4.1	3117	29	RNEWACA	R. norvegicus mRNA for	6.24e-01
41	19	4.1	4870	32	XXU13854	PGEX-4T2 cloning vecto	6.24e-01
42	19	4.1	4974	32	XXU13858	PGEX-5X3 cloning vecto	6.24e-01
43	19	4.1	4985	32	CVU78873	PGEX-6P-2 cloning vect	6.24e-01
44	19	4.1	33830	19	CEC0462	Caenorhabditis elegans	6.24e-01
45	19	4.1	103566	27	AC003087	Human BAC clone RG119H	6.24e-01

ALIGNMENTS

RESULT	1	AB016247	2125 bp	mRNA	PRI	09-OCT-1998
LOCUS		Homo sapiens mRNA for sterol-C5-desaturase, complete cds.				
DEFINITION		AB016247				
ACCESSION		AB016247				
NID		93721881				
KEYWORDS		sterol-C5-desaturase; C5D.				
SOURCE		Homo sapiens (strain:caucasian) 9-year old female liver cDNA to mRNA.				
ORGANISM		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
		Primates; Catarrhini; Hominoidea; Hominidae; Homo.				
REFERENCE	1	(sites)				
AUTHORS		Matsushima,M., Inazawa,J., Takahashi,E., Suzumori,K. and Nakamura,Y.				
TITLE		Molecular cloning and mapping of a human cDNA (SC5DL) encoding a protein homologous to fungal sterol-C5-desaturase				
JOURNAL		Cytogenet. Cell Genet. 74 (4), 252-254 (1996)				
MEDLINE		97130614				
REFERENCE	2	(sites)				
AUTHORS		Nishi,S., Nishino,H. and Ishibashi,T.				
TITLE		Molecular cloning and expression of the human and mouse lathosterol 5-desaturase				
JOURNAL		Unpublished (1998)				
REFERENCE	3	(bases 1 to 2125)				
AUTHORS		Nishino,H.				
TITLE		Direct Submission				
JOURNAL		Submitted (09-JUL-1998) to the DBJ/EMBL/GenBank databases. Hideok:				
		Nishino, Hokkaido University School of Medicine, Department of Biochemistry, N15W7 Kita-ku, Sapporo, Hokkaido 060-8638, Japan				
		(E-mail:hideakin@med.hokudai.ac.jp, Tel:+81-11-706-5047, Fax:+81-11-706-5169)				
FEATURES		Location/Qualifiers				
source		1..2125				
		/organism="Homo sapiens"				
		/strain="caucasian"				
		/db_xref="taxon:9606"				
		/dev_stage="9-year old"				

```
/sex="female"
/tissue_type="liver"
82..981
/gene="C5D"
82..981
/gene="C5D"
/EC_number="1.3.3.2"
/codon_start=1
/product="sterol-C5-desaturase"
/db_xref="PID:d1034698"
/db_xref="PID:g3721882"
/translation="MDLVLRVADYFFTPVYPVATWPEDDIFRQALSLILVTNNGAYI
LYFCFAPLSYFVFDHALMKPQFLKNQVREIKFTVQALPWLISILVALFLEIRGY
SKLHDDGEPFYLGLFELVSVIISFLFTFMFIYHRLHRLVYRLKHPHRIWIP
TPRASHAFHIDGFLQSLPYHVPFIFPLKVVVLSLYILVNIWTISIHGDGDFRVPQI
LQPPINGSAHHTDHHFEDYNGQYFTLWDRIQSGSPKNSPFFGKGPLSVVKEMTGK
RSPSPGNGCKNEKLFNGEFTKE"
725..936
/gene="C5p"
/notes="deleted in D8518 (Matsushima, M. et al. Cytogenet.
Cell Genet. 74(4), 252-254 (1996))"
/citation=[1]

BASE COUNT      670 a   388 c   379 g   588 t
ORIGIN

Query Match      6.5%; Score 30; DB 28; Length 2125;
Best Local Similarity 100.0%; Pred. No. 6.08e-11;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 15 CAGGAATCCCGGTCGACCCACGCGTCG 44
|||||
Qy 27 caggaattcccggtcgaccacgcgctcg 56

RESULT 2
LOCUS      HSRSTS367      367 bp      RNA      STS      16-SEP-1997
DEFINITION H.sapiens STS, 367bp, sequence tagged site.
ACCESSION  Y07966
NID        g1552329
KEYWORDS   STS; sequence tagged site.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 367)
AUTHORS    McKie,A.B., Iwamura,T., Leung,H.Y., Hollingsworth,M.A. and
Lemoine,N.R.
TITLE      Alu-polymerase chain reaction genomic fingerprinting technique
identifies multiple genetic loci associated with pancreatic
tumorigenesis
JOURNAL    Genes Chromosomes Cancer 18 (1), 30-41 (1997)
MEDLINE    97147122
REFERENCE  2 (bases 1 to 367)
AUTHORS    McKie,A.B.
TITLE      Direct Submission
JOURNAL    Submitted (12-SEP-1996) A.B. McKie, ICRF Molecular Pathology Group,
8th Floor, WRC Cyclotron Bldg, Hammersmith Hosp, Ducane Road,
London, W12 0NN, UK
FEATURES   Location/Qualifiers
source     1..367
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /dev_stage="fetal"
            /tissue_type="brain"
            /lab_host="E.coli"
            /chromosome="17"
            /map="p13.3"
            1..>367
BASE COUNT      61 a   146 c   95 g   65 t
ORIGIN

Query Match      6.0%; Score 28; DB 31; Length 367;
Best Local Similarity 100.0%; Pred. No. 5.65e-09;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 GGAATTCCTCCGGTCGACCCACGCGTCG 30
|||||
Qy 29 ggaattcccggtcgaccacgcgctcg 56

RESULT 4
LOCUS      HSU77594      770 bp      mRNA      PRI      25-FEB-1997
DEFINITION Human tazarotene-induced gene 2 (TIG2) mRNA, complete cds.
ACCESSION  U77594
NID        g1848263
KEYWORDS   human.
SOURCE     Homo sapiens
ORGANISM   Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 42 GGAATTCCTCCGGTCGACCCACGCGTCG 69
|||||
Qy 29 ggaattcccggtcgaccacgcgctcg 56

RESULT 3
LOCUS      HSXS16      372 bp      RNA      PRI      10-APR-1997
DEFINITION H.sapiens (xs16) mRNA, 372bp.
ACCESSION  Z36812
NID        9533926
KEYWORDS   human.
ORGANISM   Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea;
Homo.
REFERENCE  1 (bases 1 to 372)
AUTHORS    Mueller-Pillasch,F., Gress,T., Lehrach,H. and Adler,G.
TITLE      Differential gene expression in pancreatic cancer. Use of an
automated approach for the large scale isolation and
characterisation of cDNA clones containing differentially expressed
sequences
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 372)
AUTHORS    Gress,T.
TITLE      Direct Submission
JOURNAL    Submitted (16-AUG-1994) Gress T., University of Ulm, Department of
Internal Medicine I, Robert Koch Str.8, 89081 Ulm, Germany, 89081
REFERENCE  3 (bases 1 to 372)
AUTHORS    Gress,T.M., Muller-Pillasch,F., Geng,M., Zimmerhackl,F.,
Zehetner,G., Friess,H., Buchler,M., Adler,G. and Lehrach,H.
TITLE      A pancreatic cancer-specific expression profile
Oncogene 13 (8), 1819-1830 (1996)
JOURNAL    97050791
FEATURES   Location/Qualifiers
source     1..372
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="xs16"
            /dev_stage="adult"
            /tissue_type="pancreas"
            /cell_type="ductal adenocarcinoma"
            /cell_line="PATU 8988s"
            /clone_lib="xs cdna lib, pancreatic cancer cell line PATU
8988s"
            misc_feature 141..372
            /note="88% homology to bovine NADH-ubiquinone
oxidoreductase, GenEMBL-locus gb_om:btubre49 residues
78-307"
BASE COUNT      75 a   107 c   115 g   75 t
ORIGIN

Query Match      6.0%; Score 28; DB 27; Length 372;
Best Local Similarity 100.0%; Pred. No. 5.65e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 GGAATTCCTCCGGTCGACCCACGCGTCG 30
|||||
Qy 29 ggaattcccggtcgaccacgcgctcg 56

RESULT 4
LOCUS      HSU77594      770 bp      mRNA      PRI      25-FEB-1997
DEFINITION Human tazarotene-induced gene 2 (TIG2) mRNA, complete cds.
ACCESSION  U77594
NID        g1848263
KEYWORDS   human.
SOURCE     Homo sapiens
ORGANISM   Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
```

```

REFERENCE 1 (bases 1 to 770)
AUTHORS Nagpat,S., Patel,S., Jacobs,H., Disepio,D., Ghosn,C., Malhotra,M.,
Teng,M., Duvic,M. and Chandraratna,R.A.S.
TITLE Tazarotene-induced gene 2 (TIG2), a novel retinoid responsive gene
in skin
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 770)
AUTHORS Nagpat,S., Patel,S. and Chandraratna,R.A.S.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-1996) Biochemistry, Allergan Inc., 2525 Dupont
Dr., Irvine, CA 92713, USA
FEATURES
source
1..770
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="retinoid (tazarotene) treated and untreated
skin graft cultures of foreskin keratinocytes and dermal
fibroblasts"
97..588
/gene="TIG2"
97..588
/gene="TIG2"
/codon_start=1
/product="tazarotene-induced gene 2"
/db_xref="PID:g1848264"
/translation="MRRLLPLALWLVGVGVAELTEAQRRLQVLALEEFKHPVQ
WAFQTSVESAVDTFFPAGIFVRLFKQTSCKRRDKWCKPEKVRPNRKRKLACI
KLGSDEKVLGRVHCPIETQVLRAEAEHQETCLRVORAGEDPHSFYPPGQAFSKAL
PRS"
BASE COUNT 207 a 216 c 238 g 109 t
ORIGIN
Query Match 6.0%; Score 28; DB 27; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.65e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 GGAATTCCTCCGGTGCACCCACGCTCCG 36
|||||
QY 29 ggaattcccggtcgacccacgcgtccg 56

RESULT 5
LOCUS AB010946 1009 bp mRNA PLN 11-FEB-1998
DEFINITION Arabidopsis thaliana mRNA for Atrr1b, complete cds.
ACCESSION AB010946
NID 92865176
KEYWORDS Atrr1b.
SOURCE Arabidopsis thaliana (variety:columbia) cDNA to mRNA,
clone_lib:Lambda-PRL2 clone:143C4T7.
ORGANISM
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta;
Eudicotyledons; Rosidae; Capparales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Newman,T., de Bruijn,F.J., Green,P., Keeistra,K., Kende,H.,
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
Retzel,E. and Somerville,C.
TITLE Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106 (4), 1241-1255 (1994)
95148729
REFERENCE 2 (sites)
AUTHORS Sato,K., Ueda,T. and Nakano,A.
TITLE Identification of Arabidopsis thaliana RER1 homologues
Unpublished (1998)
JOURNAL 3 (bases 1 to 1009)
REFERENCE Sato,K. and Nakano,A.
AUTHORS Direct Submission
TITLE Submitted (04-FEB-1998) to the DDBJ/EMBL/GenBank databases. Ken
Sato, The Institute of Physical and Chemical Research(RIKEN),
Molecular Membrane Biology Laboratory; 2-1 Hirosawa, Wako, Saitama
351-0106, Japan [E-mail:satoken@postman.riken.go.jp,
Tel:+81-48-467-9548, Fax:+81-48-462-4679)

```

```

FEATURES
source
1..1009
Location/Qualifiers
/organism="Arabidopsis thaliana"
/variety="columbia"
/db_xref="taxon:3702"
/clone="143C4T7"
/clone_lib="Lambda-PRL2"
152..739
/gene="Atrr1b"
152..739
/gene="Atrr1b"
/codon_start=1
/product="Atrr1b"
/db_xref="PID:g2865177"
/translation="MEGSGDGSMTVPQKVVHEAWRVQYVYLDKTTPHSTNRWIGT
LVFLIYCLRVYSIHGFYIIISYGLGIVLLNLLIGFLSPVDPPELVSDGATLPTRGSD
EFKPIRLPEKFWYSMTKAFCLAFMTFESVDFVFWFNPILLCYWVFLVLTMRRO
IAHMIKHKIIPFSIGKQKYSGRKSSANGSGSRAD"
BASE COUNT 253 a 226 c 212 g 318 t
ORIGIN
Query Match 6.0%; Score 28; DB 24; Length 1009;
Best Local Similarity 100.0%; Pred. No. 5.65e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGAATTCCTCCGGTGCACCCACGCTCCG 28
|||||
QY 29 ggaattcccggtcgacccacgcgtccg 56

RESULT 6
LOCUS AF019249 1214 bp mRNA ROD 22-SEP-1997
DEFINITION Mus musculus Nmi mRNA, complete cds.
ACCESSION AF019249
NID 92425059
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Mehtani,S. and Zervos,A.S.
TITLE Mouse homologue of the human Nmi cDNA
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1214)
AUTHORS Mehtani,S. and Zervos,A.S.
TITLE Direct Submission
JOURNAL Submitted (13-AUG-1997) Cutaneous Biology Research Center,
Massachusetts General Hospital, Harvard Medical School, 13th
Street, Bldg. 149, Charlestown, MA 02129, USA
FEATURES
source
1..1214
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
150..1094
/note="Similar to H. sapiens Hou encoded by GenBank
Accession Number U32849"
/codon_start=1
/product="Nmi"
/db_xref="PID:g2425060"
/translation="MDADKDKIKOACDERSAEMDDMRGEQSGMLVHETMSKELDEE
IKKLEAEQSDAREFQIKENYPERKLLTSVSPDGDGCHFSNSCSFOVSSQILYELQ
EGQALITFEREVAQNVISMGNVHVVQMEGTPVKVSAHPVPLNTGVRFOVHVDISKMI
NVTGIPDELSEETRDKLELSEKSRNGGGEVSDYDRKRSVITFETGVGVDKIL
KKTYPLYNQKCHSVAPCIERLKYQVFSVSKKTVLLTGLLEGIPVDEETGEDL
LNIFORKNNGGGEVGVKCSLDOSFAFYFREARETI"
BASE COUNT 395 a 244 c 314 g 261 t
ORIGIN
Query Match 6.0%; Score 28; DB 29; Length 1214;
Best Local Similarity 100.0%; Pred. No. 5.65e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

九

JOURNAL MEDLINE
 REFERENCE 2 (bases 1 to 1493)
 AUTHORS Meissner,R., Jin,H.-L., Martin,C. and Bevan,M.
 TITLE Direct Submission
 JOURNAL Submitted (04-MAY-1998) Abt. Biochemie, MPI fur Zuchtungsforschung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany

FEATURES

source
 1. .1493
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="v"
 /map="near marker ve033"
 1. .1493
 /gene="MYB92"
 /allele="Col"
 260. .1264
 /gene="MYB92"
 /note="MYB92: R2R3-MYB factor family member"
 /codon_start=1
 /product="putative transcription factor"
 /db_xref="PID:g941524"
 /translation="MGKSPISDDSGLLKGGKWPDPDEKLVYVQKHGHSWRALPKLA
 GLNRCKGKSLRWTLRPDIKKRFSPEDEQITLNLHSLVIGNKWTIANQLPGRDTN
 EIRFNWHLKRLQMGDPMTHRPTDIFSLQSLMSLNLGFDVLOQQFPIDQ
 EHTILKLOTMAKLQFOYLLOPSSMNNVNPDEFTLSLLNSIASEFKETSNNTSN
 LDGLGSLYLODFSLPSLKTLSNMEPSSVFPQNLDDHFKETSTORENLVPSPILS
 DPSTTPAHVNDLLIFNQYGIENVNITSSSGESGASASAAWPDHLLDDSIIFSIP
 ."

BASE COUNT 470 a 321 c 290 g 412 t
 ORIGIN

Query Match 6.0%; Score 28; DB 25; Length 1493;
 Best Local Similarity 100.0%; Pred. No. 5,65e-09;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 25 GGAATTCGCGGTGCACCCACGCGTCGC 52
 |||||
 QY 29 ggaattccgggtgcacccacgcgcgcg 56

RESULT 10
 LOCUS I60018 1581 bp DNA PAT 15-AUG-1997
 DEFINITION Sequence 1 from patent US 5654495.
 ACCESSION I60018
 NID 92478650
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1581)
 AUTHORS Voelker,F.Alois and Davies,H.Maelor.
 TITLE Production of myristate in plant cells
 JOURNAL Patent: US 5654495-A 1 05-AUG-1997;
 FEATURES Location/Qualifiers
 source
 1. .1581
 /organism="unknown"
 BASE COUNT 418 a 336 c 410 g 397 t
 ORIGIN

Query Match 6.0%; Score 28; DB 22; Length 1581;
 Best Local Similarity 100.0%; Pred. No. 5,65e-09;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 54 GGAATTCGCGGTGCACCCACGCGTCGC 81
 |||||
 QY 29 ggaattccgggtgcacccacgcgcgcg 56

RESULT 11
 LOCUS AB000453 1640 bp mRNA PLN 28-JAN-1997

DEFINITION
 ACCESSION
 NID
 KEYWORDS
 SOURCE

ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

FEATURES

source
 1. .1640
 /organism="Petunia x hybrida"
 /cultivar="Mitchelli diploid"
 /db_xref="taxon:4102"
 /clone="ZPT3-1"
 /lab_host="Escherichia coli, DH10B"
 /tissue_type="stamen"
 95. .1408
 /note="Cys(2) His(2) zinc finger protein, 3 fingers"
 /codon_start=1
 /product="PETHY:ZPT3-1"
 /db_xref="PID:d1019856"
 /db_xref="PID:g1786138"
 /translation="MVDNSOKNEPSTVIHYCRVKRGSAGALGGMHSHGVGDHKNK
 NYGEDINEORYMINFRDKPEGQKHSYNLRANTNLLNRASEDGDKSSWPPNDR
 GYALDETLTSSMSPGSGSDLESTKPYDAKEYNGNDKXIASREEDLANCLVM
 LSKNSYLDNNNEATYKAEVEKGMFQCKACKKVFSSHQALGHRASHKVKGCYAAK
 IKDDNDNDNDNDNDNDIDEDSISPSDLIFHOESNFSQSSSSFSKRSRV
 HCSICHRVSSGQALGCHKRWLSLSPENTFTPTFOEIQYHTQEOGLFNKPMFTN
 FQPLDLNFPALGNPAEFLKLNPFHEHGRPSYVLQTLTDOQINTNLHQNEKCKDST
 EDLRREENYKDKAEKLSNLKDVNLGGSSWLVQVIGTPTDIVATL"

BASE COUNT 579 a 300 c 327 g 434 t
 ORIGIN

Query Match 6.0%; Score 28; DB 24; Length 1640;
 Best Local Similarity 100.0%; Pred. No. 5,65e-09;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 GGAATTCGCGGTGCACCCACGCGTCGC 36
 |||||
 QY 29 ggaattccgggtgcacccacgcgcgcg 56

RESULT 12
 LOCUS AF031814 1709 bp mRNA ROD 08-FEB-1998
 DEFINITION Mus musculus pregnane x receptor mRNA, complete cds.
 ACCESSION AF031814
 NID 92852328
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1709)
 AUTHORS Klierer,S.A., Moore,J.T., Wade,L., Staudinger,J.L., Watson,M.A.,
 Jones,S.A., McKee,D.D., Oliver,B.B., Willson,T.M.,
 Zetterstrom,R.H., Perlmann,T. and Lehmann,J.M.

TITLE An orphan nuclear receptor activated by pregnanes defines a novel

Petunia hybrida mRNA for PETHY:ZPT3-1, complete cds.
 AB000453
 g1786137
 zinc finger protein: ZPT3; transcription factor.
 Petunia hybrida (cultivar:Mitchelli diploid, lab_host:Escherichia
 coli, DH10B) stamen cDNA to mRNA, clone:ZPT3-1.
 Petunia x hybrida
 Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
 Charophyta/Embryophyta group; Embryophyta; vascular plants; seed
 plants; Magnoliophyta; Magnoliopsida; Solanaceae; Solanales;
 Solanaceae; Petunia.

1 (sites)
 Kobayashi,A., Sakamoto,A., Kubo,K., Rybka,Z., Kanno,Y. and
 Takatsuji,H.
 Seven zinc-finger transcription factors are expressed sequentially
 during the development of anthers and pollen in Petunia
 Unpublished (1997)
 2 (bases 1 to 1640)
 Takatsuji,H.
 Direct Submission
 Submitted (14-JAN-1997) to the DDBJ/EMBL/GenBank databases. Hiroshi
 Takatsuji, National Institute of Agrobiological Resources, Plant
 Physiology, 2-1-2 Kannondai, Tsukuba, Ibaraki 305, Japan
 (tel:0298-38-9383, fax:0298-38-7417)

Location/Qualifiers
 1. .1640
 /organism="Petunia x hybrida"
 /cultivar="Mitchelli diploid"
 /db_xref="taxon:4102"
 /clone="ZPT3-1"
 /lab_host="Escherichia coli, DH10B"
 /tissue_type="stamen"
 95. .1408
 /note="Cys(2) His(2) zinc finger protein, 3 fingers"
 /codon_start=1
 /product="PETHY:ZPT3-1"
 /db_xref="PID:d1019856"
 /db_xref="PID:g1786138"
 /translation="MVDNSOKNEPSTVIHYCRVKRGSAGALGGMHSHGVGDHKNK
 NYGEDINEORYMINFRDKPEGQKHSYNLRANTNLLNRASEDGDKSSWPPNDR
 GYALDETLTSSMSPGSGSDLESTKPYDAKEYNGNDKXIASREEDLANCLVM
 LSKNSYLDNNNEATYKAEVEKGMFQCKACKKVFSSHQALGHRASHKVKGCYAAK
 IKDDNDNDNDNDNDIDEDSISPSDLIFHOESNFSQSSSSFSKRSRV
 HCSICHRVSSGQALGCHKRWLSLSPENTFTPTFOEIQYHTQEOGLFNKPMFTN
 FQPLDLNFPALGNPAEFLKLNPFHEHGRPSYVLQTLTDOQINTNLHQNEKCKDST
 EDLRREENYKDKAEKLSNLKDVNLGGSSWLVQVIGTPTDIVATL"

BASE COUNT 579 a 300 c 327 g 434 t
 ORIGIN

Query Match 6.0%; Score 28; DB 24; Length 1640;
 Best Local Similarity 100.0%; Pred. No. 5,65e-09;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 GGAATTCGCGGTGCACCCACGCGTCGC 36
 |||||
 QY 29 ggaattccgggtgcacccacgcgcgcg 56

RESULT 12
 LOCUS AF031814 1709 bp mRNA ROD 08-FEB-1998
 DEFINITION Mus musculus pregnane x receptor mRNA, complete cds.
 ACCESSION AF031814
 NID 92852328
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1709)
 AUTHORS Klierer,S.A., Moore,J.T., Wade,L., Staudinger,J.L., Watson,M.A.,
 Jones,S.A., McKee,D.D., Oliver,B.B., Willson,T.M.,
 Zetterstrom,R.H., Perlmann,T. and Lehmann,J.M.

TITLE An orphan nuclear receptor activated by pregnanes defines a novel

```

steroid signaling pathway
Cell 92 (1), 73-82 (1998)
98149345
REFERENCE
AUTHORS
Kliwer, S.A., Moore, J.T., Wade, L., Staudinger, J.L., Watson, M.A.,
Jones, S.A., McKee, D.D., Oliver, B.B., Willson, T.M.,
Zetterstrom, R.H., Perlmann, T. and Lehmann, J.M.
Direct Submission
Submitted (29-OCT-1997) Department of Molecular Endocrinology,
Glaxo Wellcome, 5 Moore Drive, RTP, NC 27709, USA
FEATURES
Source
1. .1709
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="liver"
151. .1446
/function="steroid signaling"
/feature="pregnane-activated nuclear receptor"
/codon_start=1
/product="pregnane X receptor"
/db_xref="pid:g2852329"
/translation="MRPESRVLVCEADSALEPPIINVEEDGGLOICVCGDK
ANGYHENVTCGCKGFFRANKRNVRRCFRGTCETITRTRQOACRLKCL
GKMEMIMSDAQRALIRKREKIEAPPGQGLTEEQALIOELMDAQOTFD
TTFSHDFRFPVAFHSCPELFQASLEDPAWQIMKDRVPMKISQIRGEDS
IWNQPPSKSDGKEIPIPLHLADYSTYMFKNFAKVISYFRDLPTIEDQISLKG
TFEMLIRNFMFTETCTWECGLAYCFEPNGFGFKLLDPLMKFCHMLKQLKH
EYVLMQAISSDPDRCGVORSVVDQLOFERFALTAKYIECSRYPFAHRELFKIMA
VLTLSRNAQQTQLRIQDSHFAPLMQELFSSIDG"
BASE COUNT 449 a 444 c 468 g 348 t
ORIGIN

Query Match 6.0%; Score 28; DB 29; Length 1709;
Best Local Similarity 100.0%; Pred. No. 5.65e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 GGAATTCGGGTGACCCACGCGTCCG 35
|||||
Qy 29 ggaattccgggtgacccacgcgctccg 56

RESULT 13
LOCUS AF071316 1779 bp mRNA 03-SEP-1998
DEFINITION Mus musculus COP9 complex subunit 7a (COPS7a) mRNA, complete cds.
ACCESSION AF071316
NID 93309173
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Wei, N., Tsuge, T., Serino, G., Dohmae, N., Takio, K., Matsui, M. and
Deng, X.-W.
TITLE The COP9 complex is conserved between plants and mammals and is
related to the 26S proteasome regulatory complex
JOURNAL Curr. Biol. 8 (16), 919-922 (1998)
MEDLINE 98372877
REFERENCE
AUTHORS Wei, N., Tsuge, T., Dohmae, N., Takio, K., Matsui, M. and Deng, X.-W.
TITLE Direct Submission
Submitted (11-JUN-1998) Department of Molecular, Cellular, and
Developmental Biology, Yale University, 165 Prospect Street, New
Haven, CT 06511, USA
FEATURES
Source
1. .1779
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
1. .1779
/gene="COPS7a"
159. .986
/gene="COPS7a"

steroid signaling pathway
Cell 92 (1), 73-82 (1998)
98149345
REFERENCE
AUTHORS
Kliwer, S.A., Moore, J.T., Wade, L., Staudinger, J.L., Watson, M.A.,
Jones, S.A., McKee, D.D., Oliver, B.B., Willson, T.M.,
Zetterstrom, R.H., Perlmann, T. and Lehmann, J.M.
Direct Submission
Submitted (29-OCT-1997) Department of Molecular Endocrinology,
Glaxo Wellcome, 5 Moore Drive, RTP, NC 27709, USA
FEATURES
Source
1. .1709
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="liver"
151. .1446
/function="steroid signaling"
/feature="pregnane-activated nuclear receptor"
/codon_start=1
/product="pregnane X receptor"
/db_xref="pid:g2852329"
/translation="MRPESRVLVCEADSALEPPIINVEEDGGLOICVCGDK
ANGYHENVTCGCKGFFRANKRNVRRCFRGTCETITRTRQOACRLKCL
GKMEMIMSDAQRALIRKREKIEAPPGQGLTEEQALIOELMDAQOTFD
TTFSHDFRFPVAFHSCPELFQASLEDPAWQIMKDRVPMKISQIRGEDS
IWNQPPSKSDGKEIPIPLHLADYSTYMFKNFAKVISYFRDLPTIEDQISLKG
TFEMLIRNFMFTETCTWECGLAYCFEPNGFGFKLLDPLMKFCHMLKQLKH
EYVLMQAISSDPDRCGVORSVVDQLOFERFALTAKYIECSRYPFAHRELFKIMA
VLTLSRNAQQTQLRIQDSHFAPLMQELFSSIDG"
BASE COUNT 449 a 444 c 468 g 348 t
ORIGIN

Query Match 6.0%; Score 28; DB 29; Length 1709;
Best Local Similarity 100.0%; Pred. No. 5.65e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 GGAATTCGGGTGACCCACGCGTCCG 35
|||||
Qy 29 ggaattccgggtgacccacgcgctccg 56

RESULT 13
LOCUS AF071316 1779 bp mRNA 03-SEP-1998
DEFINITION Mus musculus COP9 complex subunit 7a (COPS7a) mRNA, complete cds.
ACCESSION AF071316
NID 93309173
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Wei, N., Tsuge, T., Serino, G., Dohmae, N., Takio, K., Matsui, M. and
Deng, X.-W.
TITLE The COP9 complex is conserved between plants and mammals and is
related to the 26S proteasome regulatory complex
JOURNAL Curr. Biol. 8 (16), 919-922 (1998)
MEDLINE 98372877
REFERENCE
AUTHORS Wei, N., Tsuge, T., Dohmae, N., Takio, K., Matsui, M. and Deng, X.-W.
TITLE Direct Submission
Submitted (11-JUN-1998) Department of Molecular, Cellular, and
Developmental Biology, Yale University, 165 Prospect Street, New
Haven, CT 06511, USA
FEATURES
Source
1. .1779
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
1. .1779
/gene="COPS7a"
159. .986
/gene="COPS7a"

/feature="pregnane-activated nuclear receptor"
/codon_start=1
/product="pregnane X receptor"
/db_xref="pid:g2852329"
/translation="MRPESRVLVCEADSALEPPIINVEEDGGLOICVCGDK
ANGYHENVTCGCKGFFRANKRNVRRCFRGTCETITRTRQOACRLKCL
GKMEMIMSDAQRALIRKREKIEAPPGQGLTEEQALIOELMDAQOTFD
TTFSHDFRFPVAFHSCPELFQASLEDPAWQIMKDRVPMKISQIRGEDS
IWNQPPSKSDGKEIPIPLHLADYSTYMFKNFAKVISYFRDLPTIEDQISLKG
TFEMLIRNFMFTETCTWECGLAYCFEPNGFGFKLLDPLMKFCHMLKQLKH
EYVLMQAISSDPDRCGVORSVVDQLOFERFALTAKYIECSRYPFAHRELFKIMA
VLTLSRNAQQTQLRIQDSHFAPLMQELFSSIDG"
BASE COUNT 449 a 444 c 468 g 348 t
ORIGIN

Query Match 6.0%; Score 28; DB 29; Length 1779;
Best Local Similarity 100.0%; Pred. No. 5.65e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 6 GGAATTCGGGTGACCCACGCGTCCG 33
|||||
Qy 29 ggaattccgggtgacccacgcgctccg 56

RESULT 14
LOCUS AF071862 2111 bp mRNA 23-JUN-1998
DEFINITION Vigna unguiculata phosphoribosylpyrophosphate amidotransferase
(pur1) mRNA, complete cds.
ACCESSION AF071862
NID 93243255
KEYWORDS
SOURCE cowpea.
ORGANISM Vigna unguiculata
Eukaryota; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta;
eudicotyledons; Rosidae; Fabales; Fabaceae; Papilionoideae; Vigna.
REFERENCE
AUTHORS Mann, A.J., Smith, P.M.C. and Atkins, C.A.
TITLE Vigna unguiculata phosphoribosylpyrophosphate amidotransferase
2 (bases 1 to 2111)
JOURNAL Unpublished
REFERENCE
AUTHORS Mann, A.J., Smith, P.M.C. and Atkins, C.A.
TITLE Direct Submission
Submitted (12-JUN-1998) Botany, University of Western Australia,
Perth, WA 6907, Australia
FEATURES
Source
1. .2111
Location/Qualifiers
/organism="Vigna unguiculata"
/strain="Vita 3"
/db_xref="taxon:3917"
/tissue_type="root nodule"
/dev_stage="day 22"
/clones="pur1-10"
1. .2111
/gene="pur1"
45. .1748
/EC_number="2.4.2.14"
/feature="PRAT"
/codon_start=1
/product="phosphoribosylpyrophosphate amidotransferase"
/db_xref="pid:g3243256"
/translation="MAASNLSTLCSSSHKSPFFNTQSFVNTNTLFNTTNSQ
TLTSHSRYTLPCTKTNTTFAFNDDEKPRECGVGVYGPDSRSLALHALQ
HRGEGAGIVADVNNVLTNGVGLVSDVDFEKLQSLPGSCAIGHVYRSTAGHRLC
NVQPFVAGYRFGSVSHVNGNFVYNSLRKLEDNGSIFNTSDTDVVLHLSKIR
PFLIYVDACENLKAGYSLVLEKLVAVKDPFGFPLVMGRSGAVFASFTCAR
DLIATYREVNPGEVVDHTGICLVTHPEPKCIFEHYFALPNSVFGSRVY
ESRKFCEILATESPVCDVIAVPDGSVVAALGYAAKAGVFOOGLIRSHYVGRFTI
EPSQKIRDFGVKLLKFLPVRGVLEKRVVVDSEIVRGTTLSKLVRLKEAGAEVHR
IACPVGSGCYGVDFESKELISNRMSVEIRFICGDSIAELPLTLRLKLEDDAP
NYCTACFSGRYIPVQPEELKTSNLNEFDWDALNGSLAKSIENEGWVRNQGVSENKIVS
V"

```

BASE COUNT 497 a 568 c 532 g 514 t
ORIGIN

Query Match 6.0%; Score 28; DB 25; Length 2111;

Best Local Similarity 100.0%; Pred. No. 5.65e-09;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 GGAATTCGGGTCGACCGCGTCCG 35

|||||

Qy 29 ggaattccgggtcgaccaccggtccg 56

RESULT 15

LOCUS AF060518 2351 bp mRNA PLN 28-OCT-1998
DEFINITION Cuphea pulcherrima 3-ketoacyl-ACP synthase (Kas4) mRNA, complete cds.

ACCESSION AF060518

NID g3800746

KEYWORDS

SOURCE

ORGANISM

Cuphea pulcherrima.

Cuphea pulcherrima

Eukaryota; Viridiplantae; Charophyta/Embryophyta group;

Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta;

Magnoliophyta; eudicotyledons; Rosidae; Myrtales; Lythraceae;

Cuphea.

1 (bases 1 to 2351)

Dehesh, K., Edwards, P., Fillatti, J., Slabaugh, M. and Byrne, J.

KAS IV: a 3-ketoacyl-ACP synthase from Cuphea sp. is a medium chain

specific condensing enzyme

Plant J. 15 (3), 383-390 (1998)

98422743

MEDLINE

REFERENCE 2 (bases 1 to 2351)

Dehesh, K.

Direct Submission

Submitted (20-APR-1998) Oils division, Calgene, 1920 Fifth Street,

Davis, CA 9561, USA

FEATURES

source

1. .2351

/organism="Cuphea pulcherrima"

/db_xref="taxon:83566"

1. .2351

/gene="Kas4"

217..1857

/gene="Kas4"

/note="KAS IV"

/codon_start=1

/product="3-ketoacyl-ACP synthase"

/db_xref="PID:g3800746"

/translation="MPAASLLASPLCTWLLACMSTSPHSDPLPPSISSPRRLSR

RRILQCPLPSASSALRSSFTLTSYACFPCDHYTISASLFGSRPIRTTRRH

RLNRASPSREANAVALQPEQVTKKPSIKQRRVVVTGMVVYPLGHPDVFYNNLL

DGTSIGSEIETFDCAQFPTRIAGEIKSFSTDGWPAPKLSKRMDFKMLYMLTAGKKALT

DGSDITDVKELDKKCGVLIGSAMGMKVNDAIEALRISVKKMNPFCVPFATTNMG

SAWLADLGMGPNYSISTACATSNFCINNAANHIIIRGEADVMLCGSDAVIPIGNG

GFVACRALSORNSDPTKAGRPWDSNDRGFVMGEGAGVLLLELEHAHKKRGATIIAEFL

GGSTFDYHMTPEHPDGAVILICIEKALAQSGVSREDVNYINAHATSTPAGDIKEYQ

ALTHCFGQNRLEKVNSTKSMIGHLLGAAGGVEAVSVQAIRTGWIHPNINLENPDGEG

DTKLLVGPKKRLNVKGLNSFGFGGHNSSILFAPYI"

BASE COUNT 613 a 563 c 537 g 537 t 1 others

ORIGIN

Query Match 6.0%; Score 28; DB 25; Length 2351;

Best Local Similarity 100.0%; Pred. No. 5.65e-09;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 23 GGAATTCGGGTCGACCGCGTCCG 50

|||||

Qy 29 ggaattccgggtcgaccaccggtccg 56

Search completed: Mon Mar 8 23:10:00 1999

Job time : 1741 secs.

THIS PAGE BLANK (USPTO)


```

QY 330 ttgccactttcgttcttatctacgtcgctg 360
      HSRSTS367      367 bp      RNA      STS      16-SEP-1997
      DEFINITION      H.sapiens STS, 367bp, sequence tagged site.
      ACCESSION      Y07986
      NID      9152329
      KEYWORDS      STS; sequence tagged site.
      SOURCE      human.
      ORGANISM      Homo sapiens
                  Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
                  Primates; Catarrhini; Hominoidea; Homo.
      REFERENCE      1 (bases 1 to 367)
      AUTHORS      McKie,A.B., Iwamura,T., Leung,H.Y., Hollingsworth,M.A. and
                  Lemoine,N.R.
      TITLE      Alu-polymerase chain reaction genomic fingerprinting technique
                  identifies multiple genetic loci associated with pancreatic
                  tumorigenesis
      JOURNAL      Genes Chromosomes Cancer 18 (1), 30-41 (1997)
      MEDLINE      97147122
      REFERENCE      2 (bases 1 to 367)
      AUTHORS      McKie,A.B.
      TITLE      Direct Submission
      JOURNAL      Submitted (12-SEP-1996) A.B. McKie, ICRF Molecular Pathology Group,
                  8th Floor, MRC Cyclotron Bldg, Hammersmith Hosp, Ducane Road,
                  London, W12 0NN, UK
      FEATURES      Location/Qualifiers
                  1..367
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /dev_stage="fetal"
                  /tissue_type="brain"
                  /lab_host="E.coli"
                  /chromosome="17"
                  /map="p13.3"
                  1..>367
      STS      61 a 146 c 95 g 65 t
      BASE COUNT
      ORIGIN
      Query Match      8.2%; Score 38; DB 31; Length 367;
      Best Local Similarity 95.1%; Pred. No. 2.19e-09;
      Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
      Db 29 CAGGTACCGTCCGGATTCGCCGGTCCGACCCACGCGTCCG 69
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      QY 16 caggtancggtcaggaattccgggtcgaccacgcgctccg 56
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      STS      61 a 146 c 95 g 65 t
      BASE COUNT
      ORIGIN
      Query Match      8.2%; Score 38; DB 31; Length 367;
      Best Local Similarity 95.1%; Pred. No. 2.19e-09;
      Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
      Db 29 CAGGTACCGTCCGGATTCGCCGGTCCGACCCACGCGTCCG 69
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      QY 16 caggtancggtcaggaattccgggtcgaccacgcgctccg 56
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      RESULT 3
      LOCUS      ATRSP40 1382 bp RNA PLN 31-OCT-1997
      DEFINITION      Arabidopsis thaliana mRNA for arginine/serine-rich splicing factor,
                  Rsp40
      ACCESSION      X99437
      NID      92582640
      KEYWORDS      splicing factor...
      SOURCE      thale cress.
      ORGANISM      Arabidopsis thaliana
                  Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
                  Embryophyta; vascular plants; seed plants; Magnoliophyta;
                  Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.
      REFERENCE      1 (bases 1 to 1382)
      AUTHORS      Lopato,S., Waigmann,E. and Barta,A.
      TITLE      Characterization of a novel arginine/serine-rich splicing factor in
                  Arabidopsis
      JOURNAL      Plant Cell 8 (12), 2255-2264 (1996)
      MEDLINE      97143875
      REFERENCE      2 (bases 1 to 1382)
      AUTHORS      Barta,A.
      TITLE      Direct Submission
      JOURNAL      Submitted (18-JUL-1996) A. Barta, Institut of Biochemistry,
                  University of Vienna, Vienna Biocenter, Dr. Bohrgasse 9/3, A-1030
                  Vienna, AUSTRIA
      Revised by [3]
      REFERENCE      3 (bases 1 to 1382)
      AUTHORS      Barta,A.
      TITLE      Direct Submission
      JOURNAL      Submitted (30-OCT-1997) A. Barta, Institut of Biochemistry,
                  University of Vienna, Vienna Biocenter, Dr. Bohrgasse 9/3, A-1030
                  Vienna, AUSTRIA
      COMMENT      On Nov 2, 1997 this sequence version replaced gi:1707367.
                  Related sequence: T20696.
      FEATURES      Location/Qualifiers
                  1..1382
                  /organism="Arabidopsis thaliana"
                  /variety="Columbia"
                  /db_xref="taxon:3702"
                  146..1198
                  /gene="RSP40"
                  146..1198
                  /gene="RSP40"
                  /codon_start=1
                  /product="splicing factor"
                  /db_xref="PID:ell69035"
                  /db_xref="PID:q2582641"
                  /translation="MKPYFCGFNDAREGDLERLFRKYKVERVDMKAGFAFVYMED
                  EDADALRALORFEGRGRRLRVEWTKSGDKRSGGSRSSMRPSKTLFVI
                  NFDADNTFRDLEKHFPYKIVNVRNFATIQYEAQDATLADANNKSLMAYK
                  ISVEYAKDDDARGHSPERRRSPSPYKRRGSPDYGRGASPLVATYK
                  ERTSPDGRRRSPSPYKRRGSPYGRDRGNDSPRRRERVASPTKYGRSPNNKPIK
                  MSPNHSPEKKSPRNGVGEVSPERRERSRSPENGQVESPGSIGRRSDGQYDCAAC
                  SPWQKSRSPRSPPADE"
      BASE COUNT      406 a 274 c 390 g 312 t
      ORIGIN
      Query Match      8.2%; Score 38; DB 24; Length 1382;
      Best Local Similarity 95.1%; Pred. No. 2.19e-09;
      Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
      Db 41 CAGGTACCGTCCGGATTCGCCGGTCCGACCCACGCGTCCG 81
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      QY 16 caggtancggtcaggaattccgggtcgaccacgcgctccg 56
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      RESULT 4
      LOCUS      AF062916 1493 bp mRNA PLN 02-DEC-1998
      DEFINITION      Arabidopsis thaliana putative transcription factor (MYB92) mRNA,
                  complete cds.
      ACCESSION      AF062916
      NID      93941523
      KEYWORDS      thale cress.
      SOURCE      Arabidopsis thaliana
                  Eukaryota; Viridiplantae; Charophyta/Embryophyta group;
                  Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta;
                  Magnoliophyta; eudicotyledons; Rosidae; Capparales; Brassicaceae;
                  Arabidopsis.
      REFERENCE      1 (bases 1 to 1493)
      AUTHORS      Kranz,H.D., Denekamp,M., Greco,R., Jin,H.-L., Leyva,A.,
                  Meissner,R., Petroni,K., Urzainqui,A., Bevan,M., Martin,C.,
                  Smeekens,S., Tonelli,C., Paz-Ares,J. and Weisshaar,B.
      TITLE      Towards functional characterisation of the members of the R2R3-MYB
                  gene family from Arabidopsis thaliana
      JOURNAL      Plant J. 16 (2), 263-276 (1998)
      MEDLINE      99056848
      REFERENCE      2 (bases 1 to 1493)
      AUTHORS      Meissner,R., Jin,H.-L., Martin,C. and Bevan,M.
      TITLE      Direct Submission
      JOURNAL      Submitted (04-MAY-1998) Abt. Biochemie, MPI fur Zuchtforschung,
                  Carl-von-Linne-Weg 10, Koeln D-50929, Germany
      FEATURES      Location/Qualifiers
                  1..1493
                  /organism="Arabidopsis thaliana"
                  /cultivar="Columbia"
                  /db_xref="taxon:3702"
                  /chromosome="V"

```

```

/imap="near marker ve033"
1. .1493
/gene="MYB92"
/allele="Col"
260. .1264
/gene="MYB92"
/note="MYB92; R2R3-MYB factor family member"
/codon_start=1
/product="putative transcription factor"
/db_xref="PID:g3941524"
/translation="MGRSPISDDSGLKKGPTWDEDEKLVNYYQKHGHSWRALPKLA
GLNRCKGRLWNTYLRPDIKRGRESPEDEQILNLHSLVGNKWSITIANOLPGRTDN
EIKNFNTHLKLLIOMGDPDTHRPRTDIFGLSOLMSLSNLRGDFVLDQQQFIDQ
EHTILKTEAKLQIFQLLPSSMSNNVNPNDFTLLSLLNSIASFKRETSNNTSN
LDLGSFTYQLQDFSLKTLNSNMEPSVFPQNDLNDHFKFSTQRENLPVSPWL
DPSSTTFAHVNDLFIQYGIEDVNSNITSSSQESGASASAAWPHLLDDSTFSDIP"
BASE COUNT 470 a 321 c 290 g 412 t
ORIGIN
Query Match 8.2%; Score 38; DB 25; Length 1493;
Best Local Similarity 95.1%; Pred. No. 2.19e-09;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 12 CAGGTACCGTCCGGATTCCCGGGTCGACCCACGCGTCCG 52
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 16 caggtancggtcaggattcccggtcgacccacgcgctccg 56

RESULT 5
LOCUS 160018 1581 bp DNA PAT 15-AUG-1997
DEFINITION Sequence 1 from patent US 5654495.
ACCESSION 160018
NID 92478650
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1581)
AUTHORS Voelker,T.Alois and Davies,H.Maelor.
TITLE Production of myristate in plant cells
JOURNAL Patent: US 5654495-A 1 05-AUG-1997;
FEATURES
source
1. .1581
/organism="unknown"
BASE COUNT 418 a 356 c 410 g 397 t
ORIGIN
Query Match 8.2%; Score 38; DB 22; Length 1581;
Best Local Similarity 95.1%; Pred. No. 2.19e-09;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 41 CAGGTACCGTCCGGATTCCCGGGTCGACCCACGCGTCCG 81
||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 16 caggtancggtcaggattcccggtcgacccacgcgctccg 56

RESULT 6
LOCUS AF060518 2351 bp mRNA PLN 28-OCT-1998
DEFINITION Cuphea pulcherrima 3-ketoacyl-ACP synthase (Kas4) mRNA, complete cds.
ACCESSION AF060518
NID g3800746
KEYWORDS
SOURCE Cuphea pulcherrima.
ORGANISM Cuphea pulcherrima
Eukaryota; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta;
Magnoliophyta; eudicotyledons; Rosidae; Myrtales; Lythraceae;
Cuphea.
REFERENCE 1 (bases 1 to 2351)
AUTHORS Dehesh,K., Edwards,P., Fillatti,J., Slabaugh,M. and Byrne,J.
TITLE KAS IV: a 3-ketoacyl-ACP synthase from Cuphea sp. is a medium chain

specific condensing enzyme
Plant J. 15 (3), 383-390 (1998)
REFERENCE 98422743
2 (bases 1 to 2351)
Dehesh,K.
Direct Submission
Submitted (20-APR-1998) Oils division, Calgene, 1920 Fifth Street,
Davis, CA 9561, USA
FEATURES
Location/Qualifiers
1. .2351
/organism="Cuphea pulcherrima"
/db_xref="taxon:83566"
1. .2351
/gene="Kas4"
217. .1857
/gene="Kas4"
/note="KAS IV"
/codon_start=1
/product="3-ketoacyl-ACP synthase"
/db_xref="PID:g3800747"
/translation="MPAASSLLASPLCTWLLAACMSTSFHPDPLPPSISSRRRLSLR
RLILSQCAPLPASASALQPEQEVTTKKPSIKORRVVTGCVVTPIGHDPDVEYNLL
RLNRASPSREMAVALQPEQEVTTKKPSIKORRVVTGCVVTPIGHDPDVEYNLL
DGTSGISLETFDCAQFPTRIAGEIKSFSTDGWAAPKLSKMDKFWMLYLTAGKALT
DGGITEDVKELDKRRKCGVLIGSAMGGMKVFNDALIEARISYKKRMDKFWMLYLTAGKALT
SAMLAMDLMGMPNYSISTACATSNFCIMNAANHIIIRGEADYMLCGSDAVIPIGMG
GFVACRALSQNRSDPTKASRPWDSNRDGFVMEGAGVLLLELEHAKRGATVIAEFL
GGFTCDAYHMTPEHPDAGVTLICIEKALQSGVSRDENVY INAHATSTPAGDIKEYQ
ALIHCGFONRELKVNSTKSMIGHLLGAGGVEAVSVQAIRTGWIHPNLENLPDEGV
DTKLLYGPKKERLNVKVLGSLNSFGGHNSSILFAPYI"
BASE COUNT 613 a 563 c 537 g 637 t 1 others
ORIGIN
Query Match 8.2%; Score 38; DB 25; Length 2351;
Best Local Similarity 95.1%; Pred. No. 2.19e-09;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 10 CAGGTACCGTCCGGATTCCCGGGTCGACCCACGCGTCCG 50
||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 16 caggtancggtcaggattcccggtcgacccacgcgctccg 56

RESULT 7
LOCUS AR024229 965 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 22 from patent US 5795961.
ACCESSION AR024229
NID 93977523
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 965)
AUTHORS Wallace,T.Paul, Harris,W.J., Carr,F.J., Old,L.J., Welt,S. and
Kitamura,K.
TITLE Recombinant human anti-Lewis b antibodies
JOURNAL Patent: US 5795961-A 22 18-AUG-1998;
FEATURES
Location/Qualifiers
1. .965
/organism="unknown"
BASE COUNT 192 a 170 c 226 g 205 t 172 others
ORIGIN
Query Match 8.0%; Score 37; DB 22; Length 965;
Best Local Similarity 20.2%; Pred. No. 1.05e-08;
Matches 22; Conservative 52; Mismatches 34; Indels 1; Gaps 1;

Db 784 HTVHVGGRVSTSTCTASDYTTSYMGVVRGWDYGGGYTNGRGRVTMDTSSNR 843
:|:::| : : :|:::| : :|:::| : :|:::| : :|:::| : :
Cp 383 ctcctgctcctccacagctcccgagctagatagaa-gcgaagatggcggaacatgg 325
:|:::| : :|:::| : :|:::| : :|:::| : :|:::| : :
Db 844 SSVTAADTAVYCVGRGRSYDSDGGDYWGDTTVTVSSHVKDMTSSSSAS 892
:|:::| : :|:::| : :|:::| : :|:::| : :|:::| : :
Cp 324 agatgaggagcgcggaatcgcggtatctatttgcctactactggagag 276
:|:::| : :|:::| : :|:::| : :|:::| : :|:::| : :
```



```

/gene="Cpl"
2310. .2426
/gene="Cpl"
/number=2
join(2328. .2426,6476. .6690,6751. .7462)
/codon_start=1
/product="cysteine proteinase-1"
/db_xref="PID:g2305221"
/translation="MRTAVLLPLALLAVAQAVSFADVYVMEWHTEKLEHRKNYODET
EERFLKIFENKHKIAKNOREAGKVSFKLAVKYADLLHHEFFKQLMNGFNVTLHK
QLRADESFGVTFISPAHVTLPKSYDWRKGAIVAKQDCHGCSWASSIGALSGQ
HFRKSGVLSQNLVDSTKYNGNCGGLMDNAPRIKNGGIDITERSYPYEAID
HFKSGKGTVDGRDFTDPOQDEGRMAEAVATGVPVSDAISHESQFYSEGVYN
EPCQDAQNLDHGLVGVFGTDESGEDYWLKNSWGTWGDGKGFIMLRNKENQCGIAS
ASSYPLV"
intron
2427. .6475
/gene="Cpl"
4346. .4553
/gene="Cpl"
/note="insertion site of P[Casper](50C)"
/citation=[1]
6476. .6690
/gene="Cpl"
/number=3
6691. .6750
/gene="Cpl"
6751. .7707
/gene="Cpl"
/number=4
join(<8110. .9300,9370. .>9532)
/product="phenylalanyl tRNA synthetase"
join(8110. .9300,9370. .>9532)
/note="potential orf"
/codon_start=1
/product="phenylalanyl tRNA synthetase"
/db_xref="PID:g2305222"
/translation="MLLTIRVOGARHLKSTRCLASAAAPKSPSPPOLEVSGSTYA
TDGWTNVTIKLSYVANKHLOTDPLSIIRQIRVNFYGVAYRNQGNPLFSYVDQMN
PVVTYQGNFDNLIIADHYVSROKSDCYIINQOHLRAHTAHQVELISGLDNFLVVG
EYVRREIDSTHVPFHQADAVRLTKDLFERNPGLFEETWSGTLPADPKLILPHP
SWTKPNPATRAVAKLMEHMKHLVGLTKDLFGPRIKIRWVDITFFPTQPSWLEEI
YFKDNWLEVGGIMRHEILQKSGVHQSIGYAFGVGLERLAWLFDIPDIRLWSDS
GFLSQFSEKDLNLPKIPISHYPOCTNDLSFWLPQDIEVDGFSNDFYDLVRVAG
DMVEQISLVDKFKPKTKGSSVCFRIVYRHMERTLQAEVNEIHKQIASASVDSFNVQ
IX"
BASE COUNT 2929 a 2357 c 2282 g 3046 t 158 others
ORIGIN

Query Match 8.0%; Score 37; DB 19; Length 10772;
Best Local Similarity 20.6%; Pred. No. 1.05e-08;
Matches 22; Conservative 55; Mismatches 28; Indels 2; Gaps 2;

Db 1691 AKRWYAWNTTMMKMWTKWMMKMTYRTWMMKMYW-TSRTTTTSAMMMYTWSTWT 1749
|: : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Cp 250 attttgaaatgaaatgcaaatcttattgaacgattatgatgatgatgatg-at 192
|: : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 1750 KYWYAWMMKMWTRTWAWAWASWAKWKTSAAAAYSARWKMKW 1796
|: : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Cp 191 gcagagagaattggattagattcagatgacagcaagcaagtcctt 145
|: : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

RESULT 10
LOCUS AC005369 74371 bp DNA PRI 01-AUG-1998
DEFINITION Homo sapiens chromosome 5, BAC clone 119j3 (LBNL H175), complete
sequence.
ACCESSION AC005369
NID g3367505
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 74371)
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.N.,
Kadner,K., Miguel,T., Miller,C., Pittluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE Sequencing of human chromosome 5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 74371)
AUTHORS Ricke,D.O.
TITLE Large Scale Sequence Analysis and Annotation with the Sequence
Comparison Analysis (SCAN) System
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 74371)
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Davis,C.A., Kadner,K., Miguel,T., Pittluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
COMMENT Sequence submitted by:
DOE Joint Genome Institute.
FEATURES
Location/Qualifiers
source
1..74371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="5q"
/clone="119j3"
/chromosome="5"
/note="LBNL H175"
893..1030
/rpt_family="Alu"
2295..2438
/rpt_family="Alu"
2818..2859
/note="(GT)21"
/rpt_type=tandem
/rpt_unit=GT
join(3246. .3410,3721. .3828)
/standard_name="RLF"
/note="65% & 69% protein identity GenPept.U22377"
3431..3724
/rpt_family="Alu"
3707..3728
/note="(A)22"
/rpt_type=tandem
/rpt_unit=A
4366..4661
/rpt_family="Alu"
5327..5602
/rpt_family="Alu"
6586..6956
/rpt_family="Li"
6847..6884
/note="(CA)19"
/rpt_type=tandem
/rpt_unit=CA
7113..7373
/rpt_family="Alu"
complement(7830..8185)
/standard_name="possible repeat"
8258..8503
/rpt_family="Alu"
9070..9387
/rpt_family="Alu"
complement(9740..9845)
/rpt_family="MER42"
complement(10440..11015)
/rpt_family="Alu"
11950..12250
/rpt_family="Alu"
12057..12085
/note="(A)29"
/rpt_type=tandem

```


